INDIAN AGRICULTURAL
RESEARCH INSTITUTE, NEW DELHI.

36765

I. A. R. I. 6.
MGIPC—S4—10 AR—21-6-49—1,000.

CHEMICAL REVIEWS

EDITORIAL BOARD
W. ALBERT NOYES, JR., Editor
LOUISE KELLEY, Assistant Editor
E. BRIGHT WILSON, JR. CARL F. CORI
N. H. FURMAN ROY C. NEWTON
M. R. FENSKE C. C. PRICE

VOLUME 44

PUBLISHED BI-MONTHLY FOR
THE AMERICAN CHEMICAL SOCIETY
BY
THE WILLIAMS & WILKINS COMPANY
Baltimore, U. S. A.
1949

CONTENTS

Number 1, February, 1949

Symposium on Thermodynamics and Molecular Structure of Solutions

Introductory Remarks. Otto Redlich 1
Equilibrium in Non-Electrolyte Mixtures. George Scatchard 7
A Critique of the Theory of Solubility of Non-Electrolytes. Joel H. Hildebrand 37
Binary Solutions of Imperfect Liquids. R. B. Parlin and H. Eyring 47
The Dielectric Constants of Hydrogen-Bonded Substances. Wendell M. Latimer 59
Critical Phenomena in Binary Liquid Systems. O. K. Rice 69
The Variation of the Properties of Electrolytic Solutions with Degrees of Dissociation. T. F. Young and L. A. Blatz 93
Activity Coefficients and Dissociation of Lead Chloride in Aqueous Solutions. Robert M. Garrels and Frank T. Gucker, Jr. 117
Aqueous Solutions of Paraffin-Chain Salts. E. C. Lingafelter 135
The Computation of the Thermodynamic Properties of Real Gases and Mixtures of Real Gases. James A. Beattie 141
Volumetric Behavior in Several Gaseous Hydrocarbon Solutions. E. W. Hough and B. H. Sage 193
The Transport Properties of Gases and Gaseous Mixtures. II. Joseph O. Hirschfelder, R. Byron Bird, and Ellen L. Spotz 205
On the Thermodynamics of Solutions. V. An Equation of State. Fugacities of Gaseous Solutions. Otto Redlich and J. N. S. Kwong 233

Number 2, April, 1949

Symposium on Aerosols

Light Scattering as a Measure of Particle Size in Aerosols. The Production of Monodisperse Aerosols. David Sinclair and Victor K. La Mer 245
Nuclei in Evaporation and Condensation. W. H. Rodebush 269
Nucleation of Supercooled Water Clouds by Silver Iodide Smokes. Bernard Vonnegut 277
The Formation of Ice Crystals in the Laboratory and the Atmosphere. Vincent J. Schaefer 291
On the Nature and Size of Particles in Haze, Fog, and Stratus of the Los Angeles Region. M. Neiburger and M. G. Wurtele 321
A Note on the Change in Average Particle Mass during the Aging of Ammonium Chloride Smokes. G. C. Langstroth, T. Gillespie, and R. M. Pearce 337
The Laws of Deposition and the Effectiveness of Insecticidal Aerosols. V. K. La Mer and Seymore Hochberg 341

The Dispersion and Deposition of Aerosols. H. F. JOHNSTONE, W. E. WINSCHE, AND L. W. SMITH 353
Electronic Methods of Counting Aerosol Particles. FRANK T. GUCKER, JR., AND CHESTER O'KONSKI 373
The Preparation and Measurement of the Concentration of Dilute Bacterial Aerosols. RONALD M. FERRY, LEO E. FARR, JR., AND MARY G. HARTMAN 389

NUMBER 3, JUNE, 1949

The Relationship between Performance and Constitution of Pure Organic Explosive Compounds. WARREN C. LOTHROP AND G. RICHARD HANDRICK 419
The Chemistry of the Oxazolines. RICHARD H. WILEY AND LEONARD L. BENNETT, JR. 447
Solubility Measurement in the Critical Region. HAROLD SIMMONS BOOTH AND RICHARD MACPHERSON BIDWELL 477

SYMPOSIUM ON THERMODYNAMICS AND MOLECULAR STRUCTURE OF SOLUTIONS[1]

Introductory Remarks

OTTO REDLICH

Shell Development Company, Emeryville, California

Received October 25, 1948

The problem of the relationship between the thermodynamic properties and the molecular structure of solutions has run through the full zigzag course which is the usual fate of a theory. The early successes of the laws of the dilute solution led to the extreme of Dolezalek's theory (10), according to which all solutions are perfect and all apparent deviations from the laws of the perfect solution are due to association of the components and to the formation of compounds.

As fashions go, the reaction came in the 1920's, under the impact of the theory of Debye and Hückel. While Dolezalek completely disregarded intermolecular and interionic forces, there appeared now a tendency to forget entirely the possible existence of chemical bonds and undissociated or associated molecules.

Actually, neither of the extreme tendencies was ever fully adopted though their influence at times was quite strong. But the proper middle course led to a difficult question: How can we give both ionic or molecular interaction and chemical bonds their due shares? Many properties, for instance, the heat of mixing, do not present a safe criterion. In extreme cases, of course, the distinction is obvious. The heat of mixing of sulfuric acid and water is undoubtedly a heat of reaction, while the heat absorbed on diluting a solution of sodium nitrate is due to interionic forces. But the heat content does not indicate a clear border line, nor does any other thermodynamic property. The same is true for other properties such as the refractivity or the viscosity.

Nevertheless a fairly clear-cut distinction, based directly on the traditional concept of the molecule, can be made. Any dissociation or association changes the number of vibrational frequencies. As soon as the vibration spectrum is fully known and interpreted, no doubt is left regarding the molecular state of a solution. But often a full knowledge of the vibrations is not necessary. A frequently discussed criterion is this: If in a solution the frequency of a Raman line or an infrared band continuously changes on changing the concentration, the shift is due to interaction with the environment. If one line gradually becomes less intense and a different line appears and becomes stronger, we may safely assume an equilibrium between two molecular species.

---

[1] Presented at the Symposium on Thermodynamics and Molecular Structure of Solutions, which was held under the auspices of the Division of Physical and Inorganic Chemistry at the 114th Meeting of the American Chemical Society, Portland, Oregon, September 13 and 14, 1948.

For an illustration of the use of vibration spectra and other properties in the elucidation of the molecular state, I should like to discuss a few examples which have found no other place on the program of this symposium.

The molecular state of nitric acid has been the subject of many investigations for a long time. Twenty years ago Hantzsch (15, 16) tried to support his theory of pseudo acids by a discussion of nitric acid, and Halban and Eisenbrand (14) encountered some difficulties in the interpretation of their careful measurements of the ultraviolet spectrum. Since that time numerous authors have contributed considerable information on the molecular state of nitric acid. Only the two latest results will be discussed here. One refers to aqueous nitric acid in the range between 80 and 100 weight per cent, the other to the range above 94 per cent and to mixtures with sulfuric acid.

According to Dalmon and Freymann (9) and to Chédin and Fénéant (5), the hydrogen band observed in the infrared and Raman spectra of aqueous nitric acid solutions undergoes a considerable change in the range between 80 and 100 per cent. The most intensive component in anhydrous nitric acid is found at 3400 cm.$^{-1}$ It gradually disappears on dilution to 80 per cent, while another intensive component at 3560 cm.$^{-1}$ and several weaker and lower frequencies appear. In the same range the deformation vibration, 608 cm.$^{-1}$, is gradually replaced by 637 cm.$^{-1}$ (5). These changes have nothing to do with ionization, since nitric acid is very little ionized in this whole range. The qualitative explanation is obvious. Anhydrous nitric acid is associated through hydrogen bonds. On addition of water, these bonds are replaced by hydrogen bonds between nitric acid and water. This process is complete at 80 per cent, which corresponds to the composition of the monohydrate. On further addition of water, ionization increases rapidly. There are various possibilities for hydrogen bonds between nitric acid molecules and between the acid and water. The available data do not appear to be sufficient for a detailed description.

For a long time it has been assumed that nitric acid in aqueous solution between 94 and 100 per cent and in mixtures with concentrated sulfuric acid is slightly decomposed into nitrogen pentoxide and water. Recently Westheimer and Kharasch (23), Bennett, Brand, and Williams (1), and Ingold and his coworkers (13, 17) have shown that a dissociation actually takes place but that it leads to a positive ion, the nitronium ion:

$$3HNO_3 = NO_2^+ + NO_3^- + HNO_3 \cdot H_2O$$
$$HNO_3 + 2H_2SO_4 = NO_2^+ + 2HSO_4^- + H_3O^+$$

The evidence is overwhelming. It includes extensive earlier experimental results of Chédin and his coworkers (4, 5, 12). A Raman line at 1394 cm.$^{-1}$ is now satisfactorily explained as due to the linear NO_2^+ ion. The freezing-point depression in sulfuric acid is in accord with the electrolytic dissociation furnishing four ions per mole of nitric acid. The absence of nitric acid lines in the Raman spectrum of solutions in sulfuric acid containing less than 15 mole per cent nitric acid and the negligible partial pressure of nitric acid over these solutions furnish further evidence. The most direct proof, however, is presented by the cathodic migration of nitrogen in sulfuric acid solutions.

Very impressive is also the influence of the nitronium equilibrium on the heat content (7). Solutions containing a little less water than corresponds to HNO_3, i.e., nominally 101 per cent HNO_3, liberate only a few hundred calories per mole of water on dilution. The heat evolution increases very rapidly and attains a maximum of 4100 cal. per mole of water at 97 per cent. On further dilution the heat effect decreases steadily. This means, of course, that the combination of nitronium ion, nitrate ion, and water to nitric acid absorbs heat, while a large amount of heat is produced by the depolymerization of the acid.

Chédin and his coworkers (4, 6) have shown, especially by the examination of Raman spectra and vapor pressures, that the molecular state of nitric acid in mixtures with sulfuric acid and water can be described in a fairly quantitative way.

The conclusions for the kinetics of nitration are obvious and well established (23): rapid nitration in anhydrous nitric acid or in mixed acid is due to the nitronium ion.

As another example I should like to discuss briefly the problem of the association of water. The basic hypothesis, that ice molecules are present in liquid water at low temperatures, was proposed by Whiting in 1884 and elaborated by Roentgen in 1892. This hypothesis was in various ways expanded by numerous authors (for earlier references see 3 and 22). The mechanism of the association was explained by Latimer and Rodebush (20) when they pointed out the existence of hydrogen bonds. The essential points of the earlier hypothesis were confirmed and elucidated by Bernal and Fowler (2; see also 21) in 1933. These authors, indeed, found by means of x-ray analysis that in cold water the structure of ice persists over short ranges (each water molecule is tetrahedrally surrounded by four molecules at a distance of 2.76 Å.). This structure disappears gradually on heating.

Qualitatively the picture is quite clear. Various numbers of water molecules are connected by hydrogen bonds to associated molecules. The equilibrium is, of course, dynamic and the life of an associative bond probably short. The larger aggregates enclose an appreciable amount of hollow space to produce the abnormally large volume of cold water. The influence of temperature, pressure, and solutes in breaking down the associated molecules is in accord with thermodynamic principles. This model explains qualitatively all the perplexing properties of water,—volume, expansibility, compressibility, specific heat, infrared and Raman spectra, viscosity, and so on.

Regarding quantitative information, we should like to know the properties and concentration of each molecular species or at least of the associative bonds. Actually our quantitative knowledge is extremely scanty. The simplest assumption would be that of a continuous association as proposed by Lassettre (19) and by Kempter and Mecke (18) for other hydroxyl compounds. But according to the results of Cross, Burnham, and Leighton (8) and of Eucken (11) it may be questioned whether continuous association represents a very good approximation. It is quite possible that some polymers, for instance $(H_2O)_8$ (11), are favored. At low temperatures there is undoubtedly a considerable amount of aggregate of higher than fourth order present, since polymers of lower order cannot enclose an appreciable hollow space. On the other hand, association to very much larger molecules is not likely because then we would have to expect a high viscosity. It was pointed out (22) in 1929 that no basis for definite statements regarding the size of the associated molecules is available. This situation has undergone little change.

Somewhat more definite is our knowledge of the heat of association. From a discussion of the Raman spectrum Cross, Burnham, and Leighton (8) derived a crude estimate of 3 kcal. for the heat of formation of a hydrogen bond. Eucken (11), on the other hand, obtained the heat of association for various polymers from the specific heat. His estimates, changed to the heat of formation of one bond, are reasonably close to 3 kcal.

Some time ago an attempt was made (22) to derive the temperature dependence of the concentration of the higher polymers from the intensity of infrared absorption. The result is in good agreement with a recent estimate of Eucken (11), based on the specific heat and the volume.

The average number of hydrogen bonds per formula weight is estimated by Cross, Burnham, and Leighton to be slightly higher than two at 26°C., according to the intensity of Raman bands.

These are meager results. But there is no reason to doubt that the quantitative description of the association equilibrium of water presents a significant and solvable problem.

Twenty years ago, when various attempts to extend the theory of Debye and Hückel to concentrated solutions were not very successful, a certain pessimism was the natural consequence. The examples discussed here indicate that the molecular theory of solutions need not be restricted to the state of high dilution.

REFERENCES

(1) Bennett, G. M., Brand, J. C. D., and Williams, G.: J. Chem. Soc. **1946,** 869, 875.
(2) Bernal, D., and Fowler, R. H.: J. Chem. Phys. **1,** 515 (1933).
(3) Chadwell, H. M.: Chem. Revs. **4,** 375 (1927).
(4) Chédin, J.: Mém. services chim. état (Paris) **31,** 113 (1944); **32,** 77 (1946).
(5) Chédin, J., and Fénéant, S.: Mém. services chim. état (Paris) **32,** 92 (1946); Compt. rend. **224,** 930, 1424 (1947).
(6) Chédin, J., Fénéant, S., and Leclerc, R.: Compt. rend. **224,** 1058 (1947).
(7) Chédin, J., Leclerc, R., and Vandoni, R.: Compt. rend. **225,** 734 (1947).
(8) Cross, P. C., Burnham, J., and Leighton, P. A.: J. Am. Chem. Soc. **59,** 1134 (1937).
(9) Dalmon, R., and Freymann, R.: Compt. rend. **211,** 472 (1940).
(10) Dolezalek, F.: Z. physik. Chem. **64,** 727 (1908); **71,** 191 (1910).
(11) Eucken, A.: Nachr. Ges. Wiss. Göttingen, Math. physik. Klasse **1946,** 38.
(12) Fénéant, S., and Chédin, J.: Compt. rend. **224,** 1008 (1947).
(13) Gillespie, R. J., Graham, J., Hughes, E. D., Ingold, C. K., and Peeling, E. R. A.: Nature **158,** 480 (1946).
(14) Halban, H. von, and Eisenbrand, J.: Z. physik. Chem. **132,** 433 (1928).
(15) Hantzsch, A.: Z. physik. Chem. **134,** 406 (1928).

(16) Hantzsch, A., and Dürigen, F.: Z. physik. Chem. **134,** 413 (1928); **136,** 1 (1928); **A144,** 153 (1929).
(17) Ingold, C. K., Millen, D. J., and Polle, H. G.: Nature **158,** 480 (1946).
(18) Kempter, H., and Mecke, R.: Z. physik. Chem. **B46,** 229 (1940).
(19) Lassettre, E. N.: J. Am. Chem. Soc. **59,** 1383 (1937); Chem. Revs. **20,** 281 (1937).
(20) Latimer, W. M., and Rodebush, W. H.: J. Am. Chem. Soc. **42,** 1419 (1920).
(21) Morgan, J., and Warren, B. E.: J. Chem. Phys. **6,** 666 (1935).
(22) Redlich, O.: Sitzber. Akad. Wiss. Wien, Math. naturw. Klasse, Abt. IIb, **138,** Suppl., 874 (1929); or Monatsh. **53/54,** 874 (1929).
(23) Westheimer, F. H., and Kharasch, M. S.: J. Am. Chem. Soc. **68,** 1871 (1946).

EQUILIBRIUM IN NON-ELECTROLYTE MIXTURES[1]

GEORGE SCATCHARD

Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts

Received October 25, 1948

The following subjects are discussed or reviewed: *(1)* the analytical expression of an extensive property interesting in thermodynamics as quantity per mole of unmixed components plus the product of two mole fractions multiplied by a power series in the difference between these mole fractions; *(2)* the related expressions for the two partial molal quantities and their difference; *(3)* the illustration of these methods applied to volume changes on mixing and to excess free energies; *(4)* the approximation of properties of polycomponent systems to those of the two-component systems comprising them; *(5)* the computation of analytical expressions for the excess free energy from measurements of vapor pressure and composition as functions of liquid composition, corrected or uncorrected for deviations of the vapor from perfect gas laws; *(6)* a "Gibbs–Duhem" test of consistency of vapor equilibrium measurements without differentiation, by plotting short segments of lines through corresponding potentials at $x = 0$ and 1, which should be tangent to the curve of molal free energy; *(7)* the computation of excess free energies from vapor pressures and liquid compositions, from liquid and vapor compositions, from the vapor pressure of one component, from azeotropic pressure and composition, from the solubility of a solid component, or from the composition of two saturated liquid phases; *(8)* the relation of the mole fraction expansion to the corresponding expression for the enthalpy of mixing in terms of volume fractions or other units, and to the Flory–Huggins expression for the excess entropy of mixing; *(9)* the possibility of expressing these two relations in terms of surface fractions rather than volume fractions; *(10)* the differences between the free energy, enthalpy, and entropy of mixing at constant pressure and the work content, energy, and entropy of mixing at constant total volume, and their effects upon the comparison of calculated and measured quantities; *(11)* the possibility of the cohesive energy density important in solutions being somewhat larger than the energy difference between the perfect gas and liquid states because of the intramolecular cohesive energy of the gas; *(12)* the molal volumes, the product of molal volume and the square root of the cohesive energy density of an homologous series and of the isomeric octanes; *(13)* the calculation of cohesive energy densities as proportional to the five-thirds power of the refraction; *(14)* the sorting or non-random distribution of molecules in solution; *(15)* partial accounting for sorting by use of the law of mass action for association of one component with one reactive group per molecule giving only dimers, with two reactive groups per molecule giving all linear polymers, with three or more reactive groups per molecule giving three-dimensional polymers, and for one-to-one combination.

INTRODUCTION

Most of our treatment of equilibrium is based upon the finding of Willard Gibbs (5) that, under certain simplifying conditions, the temperature, pressure, and chemical potential of each component are each constant throughout a system in equilibrium. In some cases some of the simplifying conditions have to be removed, but we know pretty well now how to handle these more complicated cases. As a consequence we usually take temperature and pressure as independent variables. However, we seldom take the potentials. Instead, we usually use the number of units of each component and most often we choose as unit the mole. We reduce these to intensive quantities by dividing the number of moles of each component by the total number of moles in the phase, to give the composition as mole fractions of the components. It is then convenient to reduce the other extensive quantities to intensive ones in the same way. For the study of polycomponent systems, or for the determination of one kind of equilibrium from measurements of another kind, it is essential to have analytical expressions for these functions. It is always useful to have at least an approximate analytical expression, so that the small deviations may be plotted on a large scale.

---

[1] Presented at the Symposium on Thermodynamics and Molecular Structure of Solutions, which was held under the auspices of the Division of Physical and Inorganic Chemistry at the 114th Meeting of the American Chemical Society, Portland, Oregon, September 13 and 14, 1948.

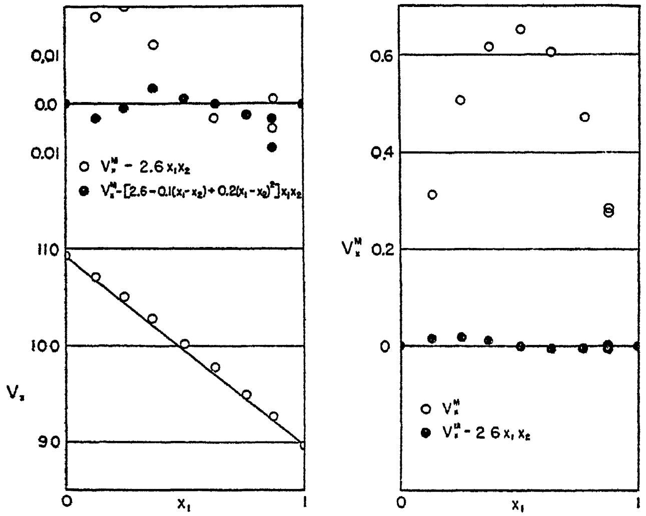


FIG. 1. Benzene–cyclohexane: volumes

VOLUME OF MIXING

We consider the volume first, because the relations are simple and the measurements are so direct that each step can be easily visualized. The phase consists of N_J moles of component J, N_K of component K, etc., and it has a volume V. Then the mole fraction x_K is defined by the relation $x_K = N_K/\Sigma_J N_J$, in which Σ_J means a sum over all the components. The volume per mole is defined by the relation, $V_x = V/\Sigma_J N_J$ and the partial molal volume of component K is defined by the relation

$$\bar{V}_K = (\partial V/\partial N_K)_{T,p,N} = (\partial V_x \Sigma_J N_J/\partial N_K)_{T,p,N}$$

In the pure component K, $\bar{V}_K = V/N_K = V_K$, which we call the molal volume of component K.

Figure 1 shows on the left V_x at 30°C. as a function of x_1 for mixtures of benzene (component 1) and cyclohexane (component 2) (21). Although in this case the change is only about 20 per cent, it is obvious that the accuracy of plotting can be increased by subtracting a straight line through the end points to give the volume of mixing per mole, V_x^M, defined by the relation

$$V_x^M = V_x - x_1 V_1 - x_2 V_2$$

This is represented in the upper scale on the right. The scale of ordinates is 50 times that of the plot of V^M, but a parabola will allow still more accurate plotting. The open circles in the upper left-hand section of figure 1 are

$$V_x^M - 2.6 x_1 x_2$$

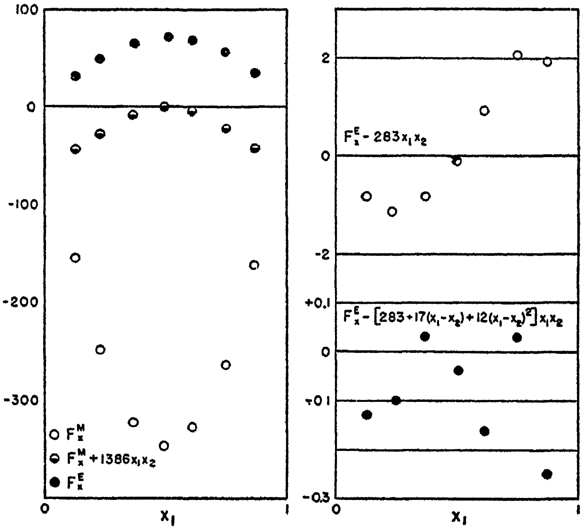


FIG. 2. Benzene–cyclohexane: free energies

The deviations are still apparently smooth, so that it should pay to consider higher terms. Since the parabola accounts for most of the volume of mixing, it is convenient to keep this term and to expand in the form

$$V_x^M = x_1 x_2 \, \Sigma_\nu V_{12}^\nu (x_1 - x_2)^\nu$$

with ν summed from 0 to n.[2] V_{12}^ν is the coefficient of the ν^{th} term (it is not V_{12} to the ν^{th} power). If we were trying to do the best job possible, we would use the method of least squares to determine the coefficients. Even then it is often

---

[2] This form was suggested by Guggenheim (6) for the excess free energy, and has since been used by many others. Redlich and Kister (15) have shown that with this form many polycomponent systems can be treated to a good approximation by adding the expressions for all the binary pairs composing them. A less elegant expression of this function was used by Scatchard and Hamer (18) to compare the development in mole fractions with that in volume fractions, and the latter expression was used for polycomponent mixtures by Scatchard, Wood, and Mochel (23).

For a binary system there is no fundamental difference between this and other methods of expressing concentrations, such as that used by Benedict, Johnson, Solomon, and Rubin convenient to work with deviation functions. Since we have already pinned down the function at the two ends and the middle, it is convenient to pin it down also at the one-quarter and three-quarters points and to determine two more constants by the relations:

$$V_{0.75} - V_{0.25} = 3V'_{12}/16$$

$$V_{0.25} + V_{0.75} = 3V^0_{12}/8 + 3V''_{12}/32$$

$V_{0.75}$ is the molal volume when $x_1 = 0.75$, and $V_{0.25}$ the volume when $x_1 = 0.25$,

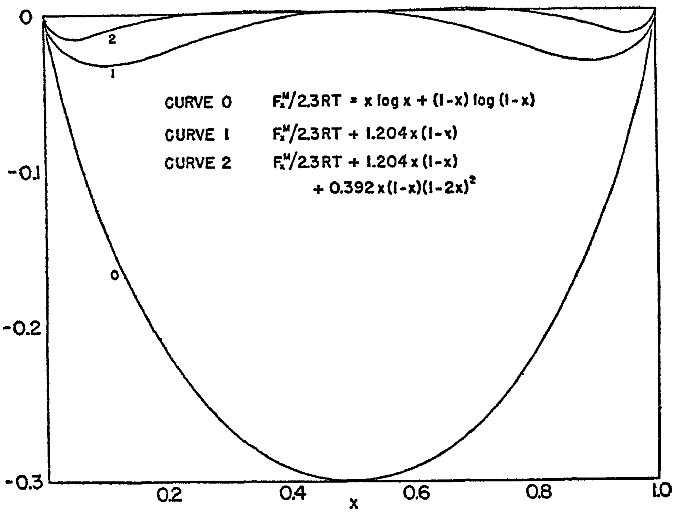


FIG 3. Ideal free energy of mixing

The filled circles in the upper left-hand section of figure 1 are

$$V^M_x - x_1x_2[2.6 - 0.1(x_1 - x_2) + 0.2(x_1 - x_2)^2]$$

The scale is 500 times that of V_x. Except for the last point the average deviation is less than 0.002 per cent.

THE EXCESS FREE ENERGY

The free energy of mixing per mole at 50°C., F^M_x, for the same system (21) is shown in the left side of figure 2 as open circles. The half-filled circles are $F^M_x + 1386x_1x_2$. The parabola does not give a satisfactory correction near the ends. If there had been accurate measurements in more dilute solutions, the discrepancy would have been even more obvious. The reason is that the slope of F^M_x is infinite at the ends, even for ideal solutions. Figure 3 shows the free energy of mixing for an ideal solution as curve 0. Curve 1 is corrected for the parabola and curve 2 for the fourth-degree term. All odd terms are zero, since the curve has axial symmetry.

---

(1) for the excess free energy of mixing. If one constant is used, their $2A_{12} = F^0_{12}$; for two constants $3A_{112} = F^0_{12} + F'_{12}$, $3A_{122} = F^0_{12} - \mathrm{F}_{12}$; and for three constants $4A_{1112} = F^0_{12} + F'_{12} + F''_{12}$, $3A_{1122} = F^0_{12} - F''_{12}$, $4A_{1222} = F^0_{12} - F'_{12} + F''_{12}$.

So we define the excess free energy of mixing per mole, F^E_x, as the free energy of mixing minus that of an ideal solution, or

$$F^E_x = F^M_x - RT\,\Sigma_j x_j \ln x_j$$

F^E_x for this system is plotted as filled circles on the left of figure 2.

The open circles on the right are

$$F^E_x - 283x_1x_2$$

on a 50 times larger scale, and the filled circles are

$$F^E_x - x_1x_2[283 + 17(x_1 - x_2) + 12(x_1 + x_2)^2]$$

on a scale 100 times larger again. The agreement could be improved still more by a more careful choice of constants.

POWER SERIES EXPANSION

These two applications show the usefulness of the method and its limitations. It can be applied to any function which is never infinite itself and has no infinite derivatives. It fails for F^M_x because of the infinite slope at zero concentration. It will fail also for F^E_x, V^M_x, or any similar functions of dilute electrolyte solutions. The Debye theory tells us that for such solutions F^E_x, V^M_x, H^M_x, etc. are proportional to the square root of the mole fraction of the ions when this mole fraction is very small. The theory also enables us to calculate an electrostatic term F^D_x, such that $(F^E_x - F^D_x)$ may be treated as F^E_x for a non-electrolyte mixture. One of the important achievements of the Debye theory is to confirm the assumption which had been made empirically for many years, that for very dilute solutions of non-electrolytes F^E_x is proportional to x. The method can be applied, of course, to the enthalpy and energy of mixing H^M_x and E^M_x, as to the volume, and to the excess entropy and work content, $S^E_x = S^M_x + R\Sigma_i x_i \ln x_i$ and $A^E_x = A^M_x - RT\Sigma_i x_i \ln x_i$, as to the excess free energy.

If a tangent is drawn to a curve of a molal quantity against the mole fraction, the intercepts at $x_1 = 1$ and at $x_2 = 1$ are the corresponding partial molal quantities of components 1 and 2, respectively. Thus when V_x is plotted against x, the intercepts of the tangent are $\bar{V}_1$ and $\bar{V}_2$; for V^M_x the intercepts are $\bar{V}^M_1 = \bar{V}_1 - V_1$ and $\bar{V}^M_2 = \bar{V}_2 - V_2$. For F^M_x the intercepts are $\mu^M_1 = RT \ln a_1$ and $\mu^M_2 = RT \ln a_2$, and for F^E_x they are $\mu^E_1 = RT \ln \gamma_1$ and $\mu^E_2 = RT \ln \gamma_2$, in which a_1 and a_2 are the activities of the components with the component liquids as the standard states, and $\gamma_1 = a_1/x_1$ and $\gamma_2 = a_2/x_2$ are the corresponding activity coefficients.

For the deviation curves it is necessary to add to the measured intercept the analytical expression corresponding to the equation used.

If

$$G_x = x_1x_2\,\Sigma_\nu G^\nu_{12}(x_1 - x_2)^\nu$$

$$\frac{dG_x}{\partial x_2} = -\frac{dG_x}{\partial x_1} = \bar{G}_2 - \bar{G}_1 = \Sigma_\nu G^\nu_{12}[(x_1 - x_2)^{\nu+1} - 2\nu x_1x_2(x_1 - x_2)^{\nu-1}]$$

$$\bar{G}_1 = x_2^2\,\Sigma_\nu G^\nu_{12}[(x_1 - x_2)^\nu + 2\nu x_1(x_1 - x_2)^{\nu-1}]$$

$$\bar{G}_2 = x_1^2\,\Sigma_\nu G^\nu_{12}[(x_1 - x_2)^\nu - 2\nu x_2(x_1 - x_2)^{\nu-1}]$$

in figure 4 are plotted the contributions of the first five terms to G_x, each expressed as G_x^ν/G_{12}^ν. Note that the maximum effect decreases as ν becomes larger, and that the maxima spread away from the center. It is obvious that $G_{21}^\nu = G_{12}^\nu(-1)^\nu$. Figure 5 shows the corresponding values of $\bar{G}_2 - \bar{G}_1$. It is worthy of note that the term with $\nu = 1$ is the only one which is not zero at the midpoint. Figure 6 gives the corresponding curves for $\bar{G}_2$. Those for $\bar{G}_1$ are obtained by interchanging subscripts and multiplying by $(-1)^\nu$. Each term beyond the first changes sign and each beyond the second changes sign twice, once at the midpoint and once at a lower concentration.

The excess free energy makes the best function for correlating vapor–liquid equilibrium data when both the equilibrium pressure and the vapor composition have been determined. Each point represents two potential measurements weighted as they would be for a least squares treatment, by multiplying by the corresponding mole fraction. It is sometimes convenient to use the corresponding function uncorrected for deviations from the perfect gas laws and divided by 2.3RT:

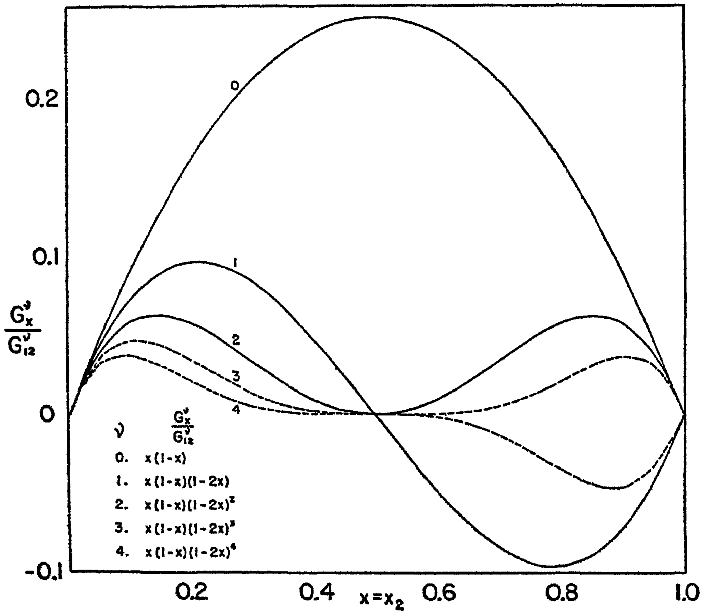


FIG. 4. Expansion of G_x

$$Q_x = x_1 \log Py_1/P_{10}x_1 + x_2 \log Py_2/P_{20}x_2$$

Figure 7 shows Q_x for chloroform–ethanol mixtures at 45°C. as open circles (20); x_2 is the mole fraction of ethanol. The filled circles are $Q_x - 0.470x_1x_2$ and the open circles near the zero line are

$$Q_x - x_1x_2[0.470 + 0.219(x_1 - x_2)] \text{ at } 45°\text{C}.$$

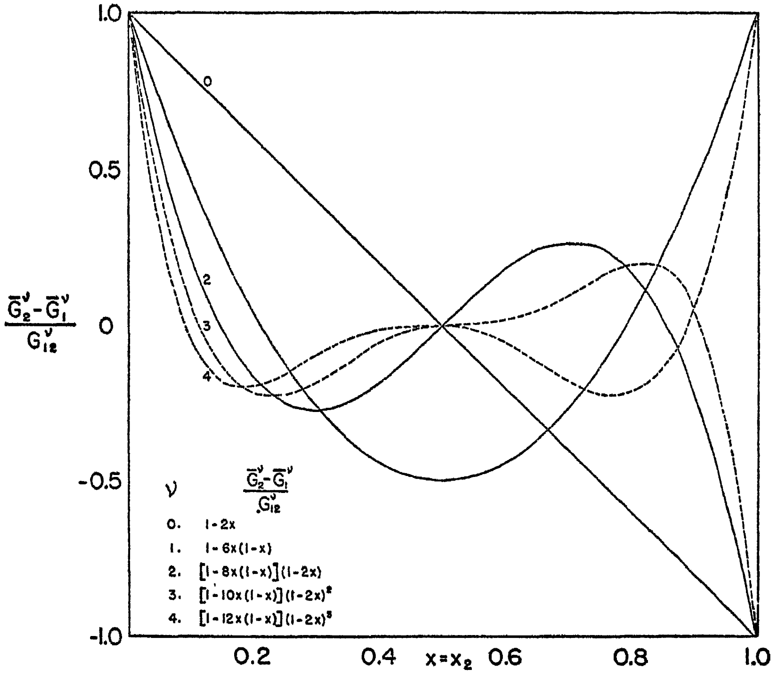


FIG. 5. Expansion of $\bar{G}_2 - \bar{G}_1$

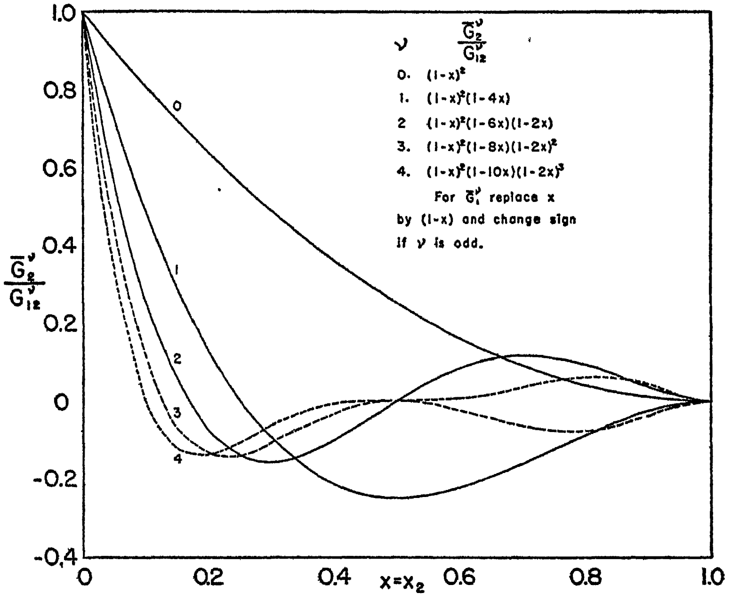


FIG. 6. Expansion of $\bar{G}_2$

These last are replotted on a larger scale in figure 8, together with the corresponding functions

$$Q_x - x_1x_2[0.462 + 0.255(x_1 - x_2)] \text{ at } 35°\text{C}.$$

and

$$Q_x - x_1x_2[0.455 + 0.183(x_1 - x_2)] \text{ at } 55°\text{C}.$$

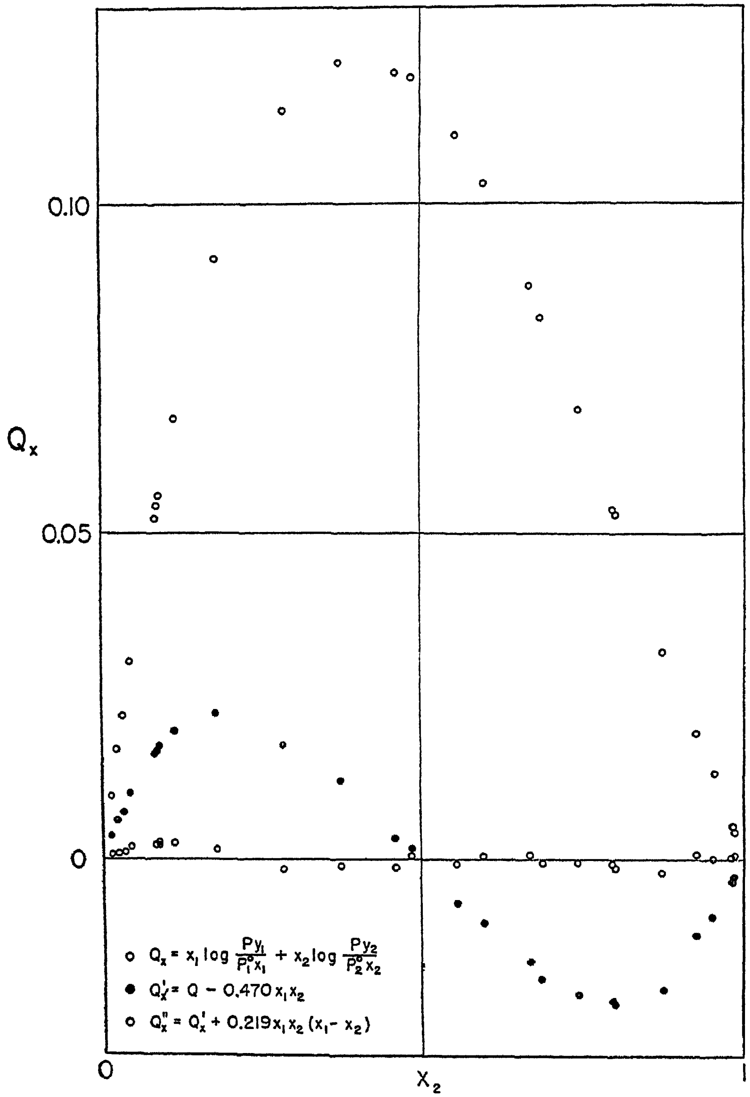


FIG. 7. Chloroform–ethanol: vapor pressures at 45°C.

The next term in the expansion with $\nu = 2$ is very small. The curves are drawn to give an approximate fit for $\nu = 3$ at $x = \frac{1}{8}$ and $\frac{7}{8}$ as well as at $\frac{1}{4}$, $\frac{1}{2}$, and $\frac{3}{4}$. The curves at all temperatures are for the function

$$0.078x_1x_2(x_1 - x_2)[(x_1 - x_2)^2 - 0.25]$$

This term gives a considerable improvement in the fit at 35° and 45°C., but not much improvement at 55°C.

This analytical expression of the partial pressures is often called an integration of the Gibbs–Duhem equation. The Gibbs–Duhem equation is not sufficient because it also applies when the slope is infinite, as in the free energy of mixing. However, the measurements do give more than the excess free energy of mixing; they give the excess potential of each component. The consistency of the results may be determined by a reversal of Gibbs's method of determining the

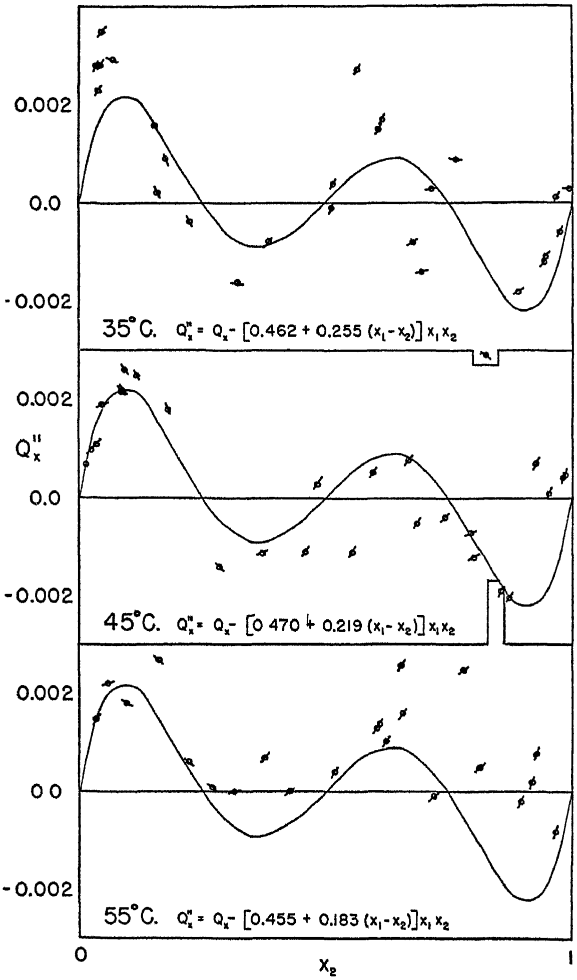


FIG. 8. Chloroform–ethanol: the curves represent $0.078x_1x_2(x_1 - x_2)[(x_1 - x_2)^2 - 0.25]$

potentials as the intercept of the tangent to the curve of G_x *vs.* x_1 with the ordinates $x_1 = 1$ and $x_2 = 1$. If the measured value of $\bar{G}_1$ at $x_1 = x$ is plotted at $x_1 = 1$, the corresponding value of $\bar{G}_2$ at $x_2 = 1$ and the straight line drawn through them, the intercept with the ordinate $x_1 = x$ is the corresponding value of G_x and the line is the tangent to the curve of G_x *vs.* x_1 at that point. For a difference function we substitute the corresponding differences in the potentials. In practice it is simpler to calculate G_x algebraically as $x_1\bar{G}_1 + x_2\bar{G}_2$ and to draw a short segment of the line through G_x and one of the end points. The segments of straight lines should be tangent to the smooth curve through all the points. These lines are drawn in figure 8.

This method has the advantage over any other "test of the Gibbs–Duhem equation" which I know in that it involves no differentiation, but the measurements at each point are compared with the smooth curve through all the points. It shares the disadvantage of all such tests that we do not know how large an error in the measurements corresponds to a given deviation in slope. We can tell that the deviation of a single point from the smooth curve through the other points probably means an error in the pressure if the slope is correct but the position is in error, and probably an error in y if the position is correct and the slope in error. Deviations in both slope and position may indicate an error in x. A more exact comparison may be obtained by calculating P and y from the equation and comparing each with the corresponding measured quantity.

The scale of figure 8 is such that it should show the deviations of Q_x from an extensive property. At 45°C. the correction may be represented quite accurately by $F_x^E/2.3RT - Q_x = -0.01x_1x_2$, but the tangent does not give the corresponding values of the potentials. It would be hard to say, however, that the corrected points are more consistent than the uncorrected ones.

If only the liquid and vapor compositions have been determined, the coefficients can be determined from the equation

$$\begin{aligned}\log \alpha_{12} &= \log y_1x_2/x_1y_2 \cong \frac{1}{2.3RT}\,\frac{\mathrm{d}F_x^E}{\mathrm{d}x_1} + \log \frac{P_{10}}{P_{20}} \\ &= A - \Sigma_\nu Q_{12}^\nu[(x_1 - x_2)^{\nu+1} - 2\nu x_1x_2(x_1 - x_2)^{\nu-1}]\end{aligned}$$

Redlich and Kister (15) have discussed the determination of the parameters of this equation. If we pin down the curve at $\frac{1}{4}$, $\frac{1}{2}$, and $\frac{3}{4}$, with the values y_1, y_2, y_3:

$$\begin{aligned}A &= 2(y_1 + y_3 - y_2/2)/3 \\ Q_{12}^0 &= y_1 - y_3 \\ Q_{12}' &= 4(2y_2 - y_1 - y_3)/3\end{aligned}$$

If only total pressures have been measured, the vapor pressures of the components and of one mixture are sufficient to give one constant. If the mixture is equimolal

$$\log P_{0.5} = \log (P_{10} + P_{20})/2 + Q_{12}^0/4$$

Further approximations must be made by trial and error. If the vapor pressures of the two components are approximately equal, two more constants can be obtained readily from the measurements at $x = \frac{1}{4}$ and $\frac{3}{4}$ in addition to the measurements used above.

If only one component is volatile, all the parameters contain the square of the mole fraction of the other component as a factor. There is therefore a great advantage in using measurements in solutions as dilute as possible. A plot of log Py/P_1x_1 minus an analytical function *vs.* x_2^2 will indicate the next function to choose.

The azeotropic pressure and composition are sufficient to determine two constants if the vapor pressures of the components at the same temperature are known. They may be determined from the equations

$$\log P_A/P_{10} = (1 - x_A)^2[Q_{12}^0 + Q'_{12}(3 - 4x_A)]$$

$$\log P_A/P_{20} = x_A^2[Q_{12}^0 + Q'_{12}(1 - 4x_A)]$$

in which P_A is the azeotropic pressure and x_A the mole fraction of component 1 in the azeotrope. The solubility of one solid component in the other will give one constant if the difference in potential of the solid and liquid is known from thermal measurements or from other solubility measurements. If the system is only partially miscible, the compositions of the two phases in equilibrium determine two constants (18).

When the deviations are large and positive it is always well to check that the equation does not demand separation into two liquid phases when there is none. This can be shown by the plot F_x^M, which shows an inflection if there is separation into two phases. Usually it is shown more clearly by plotting μ_1^M or μ_2^M, which show a maximum and minimum if there are two liquid phases and a horizontal inflection at the critical mixing point.

EFFECT OF DIFFERENCES OF VOLUME OR SURFACE

Expansion in power series of the mole fraction is not always the most economical expression of these functions. Eighteen years ago I showed (17) that for simple binary non-electrolyte mixtures, the excess free energy per unit volume should be expressed approximately by a constant times the product of the volume fractions of the components, that the constants should be calculable from the energies of evaporation of the components, and that the excess free energy of polycomponent solutions can be calculated with no new parameters. We have also shown in several papers that introduction of a second term is more effective with this expression than with mole fractions, although mixtures of a non-polar compound with an alcohol seem to be represented more simply by the mole fraction form. The simple one-parameter equation has since been much used by Hildebrand and his collaborators with solutions of small molecules and by many high-polymer chemists for the enthalpy of mixing. Professor Hildebrand discusses it more fully in this symposium (10). For a two-component system, if $\rho = V_2/V_1$,

$$F_x^E = \beta\,\frac{V_1x_1V_2x_2}{V_1x_1 + V_2x_2} = \frac{2\beta V_2x_1x_2}{\rho + 1}\;\frac{1}{1 - \dfrac{\rho - 1}{\rho + 1}\,(x_1 - x_2)}$$

$$\frac{2\beta V_2}{\rho + 1}\,x_1x_2\,\Sigma_\nu\left(\frac{\rho - 1}{\rho + 1}\right)^\nu (x_1 - x_2)^\nu$$

So

$$F_{12}^{\nu} = \frac{2\beta V_2}{\rho + 1}\left(\frac{\rho - 1}{\rho + 1}\right)^{\nu}$$

The corresponding expressions for the chemical potentials are

$$\mu_1^E = \beta V_1\left(\frac{V_2 x_2}{V_1 x_1 + V_2 x_2}\right)^2 = \frac{4\beta V_1 \rho^2 x_2^2}{(\rho + 1)^2}\left(\frac{1}{1 - \frac{\rho - 1}{\rho + 1}(x_1 - x_2)}\right)$$

$$\mu_2^E = \beta V_2\left(\frac{V_1 x_1}{V_1 x_1 + V_2 x_2}\right)^2 = \frac{4\beta V_2 x_1^2}{(\rho + 1)^2}\left(\frac{1}{1 - \frac{\rho - 1}{\rho + 1}(x_1 - x_2)}\right)$$

The series F_{12}^{ν} converges for any finite value of ρ, and it converges rapidly if ρ is near unity, say between $\frac{1}{3}$ and 3. The mole fraction form is not a useful one for high polymers because of the slowness of convergence and because it fails to show simply that the enthalpy of mixing of a given mass of polymer is nearly independent of its molecular weight.

These relations hold equally well if ρ be defined differently, provided the product βV_2 be defined to correspond. The equation of Langmuir (13) is equivalent to defining ρ as $(V_2/V_1)^{2/3}$, and that used by van Laar (25) is equivalent to defining ρ as b_2/b_1, the ratio of van der Waals b's for the two components, which is almost the same as taking V_2/V_1. In the "van Laar equation" often used by chemical engineers, however, both ρ and βV_2 are determined empirically, usually as βV_2 and $\beta V_2/\rho$.

In my first treatment I assumed that simple non-polar solutions are regular, that is, that the entropy of mixing is the same as for an ideal solution. Almost everyone treating solutions was making this assumption, and the attempts to prove it were often more amusing than convincing. I accepted Hildebrand's justification (8) that solutions are regular when "orienting and chemical effects are absent and the distributions and orientations are random."

Since then the high-polymer chemists have shown that volume differences lead to an excess entropy of mixing. Starting with a quasi-crystal lattice model in which each solvent molecule and each polymer unit occupies one unit cell, Huggins (11) and Flory (3) arrived independently at expressions which may be written as

$$S_x^E = -R\,\Sigma_j x_j \ln V_j/V_x$$

This question has been treated more elegantly by Guggenheim (7). Although the original derivations apply to linear polymers which cannot coil back on themselves, Huggins suggests that the equation should apply to any mixture. Hildebrand (9) has recently reached the same conclusion, without using a quasi-crystal lattice, but with the assumptions that there is an entropy term proportional to the "free volume," and that the "free volume" per mole of a liquid is proportional to its total volume. This second assumption does not seem very probable to me.

For a binary mixture, the Flory–Huggins expression may be written:

$$-\frac{S_x^E}{R} = x_2 \ln \rho - \ln [1 + (\rho - 1)x_2]$$

The corresponding values for the partial molal entropies are

$$-\frac{\bar{S}_1^E}{R} = \ln \frac{1}{1 + (\rho - 1)x} + \frac{(\rho - 1)x}{1 + (\rho - 1)x}$$

$$-\frac{\bar{S}_2^E}{R} = \ln \frac{\rho}{1 + (\rho - 1)x} - \frac{(\rho - 1)(1 - x)}{1 + (\rho - 1)x}$$

In comparing other analytical expressions with the power series expansion, we may use two methods, both of which seem to me dangerous, however, when applied to curves from experimental measurements. The first is to determine the coefficients of the power series from the value of the other function and its derivatives at $x = 0.5$ by the relation

$$G_{12}^{\nu} = \frac{4}{2^{\nu}\nu!}\left(\frac{\mathrm{d}^{\nu} G_x}{\mathrm{d}x^{\nu}}\right)_{0.5}$$

taking G_{12}^0 as $4\left(\frac{\mathrm{d}^0 G_x}{\mathrm{d}x_1^0}\right)_{0.5}$. The second is to utilize the extrapolation of $\bar{G}_1^E$ to $(\bar{G}_2^E)_2$ at $x_2 = 1$ and of $\bar{G}_2$ to $(\bar{G}_2^E)_1$ at $x_1 = 1$ to give:

$$[(\bar{G}_1^E)_2 + (\bar{G}_2^E)_1]/2 = \Sigma G_{12}^{\nu} \text{ even}$$

$$[(\bar{G}_2^E)_2 - (\bar{G}_1^E)_1]/2 = \Sigma G_{12}^{\nu} \text{ odd}$$

Carlson and Coburn (2) use these two equations to determine two parameters from the extrapolation of experimental measurements. In the Margules equation their parameters are F_{12}^0 and F_{12}'; in the "van Laar equation" they are βV_2 and $\beta V_1 = \beta V_2/\rho$.

It is possible to determine any number of parameters from the derivatives at the midpoint and to check by extrapolation of the partial quantities to determine the residuals such as $\Sigma G_{12\text{ even}} - G_{12}^0$ or $\Sigma G_{12\text{ even}} - G_{12}^0 - G_{12}''$ and $\Sigma G_{12\text{ odd}} - G_{12}'$.

Figure 9 shows F_{12}'/F_{12}^0, F_{12}''/F_{12}^0, $(\Sigma F_{12\text{ odd}} - F_{12}')/F_{12}^0$ and $(\Sigma F_{12\text{ even}} - F_{12}^0 - F_{12}'')/F_{12}^0$ for the volume fraction expression given above, with F_{12}' and F_{12}'' determined from the derivatives at the midpoint. The dotted lines are F_{12}'/F_{12}^0 and F_{12}''/F_{12}^0 with F_{12}' and F_{12}'' determined from the values of F_{12}^0 and F_x^E at $x = \frac{1}{4}$ and $\frac{3}{4}$. The value of F_{12}^0 does not depend upon ρ.

Figure 10 shows the first four curves for the Flory–Huggins entropy. The value of F_{12}^0 does depend upon ρ, and is given as the broken curve with ordinates at the right. The Flory–Huggins entropy is more symmetrical than the volume fraction expression for the same value of ρ. It corresponds roughly to the latter at $\rho^{2/3}$.

The Flory–Huggins treatment gives too large an effect in many cases. Zimm (28) has assumed that high-polymer molecules are so much bigger than those of the solvent that the solvent may be treated as a continuous medium, and has made calculations for dilute solutions which are equivalent to the calculation of van der Waals' b for spherical molecules and for long cylinders. His results for long cylinders are larger than the simple Flory–Huggins result. The quantity $(\pi/RTc - 1)/c$, with π the osmotic pressure, is $\bar{V}_2 - \bar{V}_1/2$ for ideal solutions. Zimm obtains $l\bar{V}_2/d$ for cylinders with length, l, greater than ten times the diameter, d. By the Flory–Huggins assumptions this is $\bar{V}_2^2/\bar{V}_1$, which is twice their value for this term. For large spheres, however, Zimm obtains $4\bar{V}_2$, which is much smaller than the Flory–Huggins value.

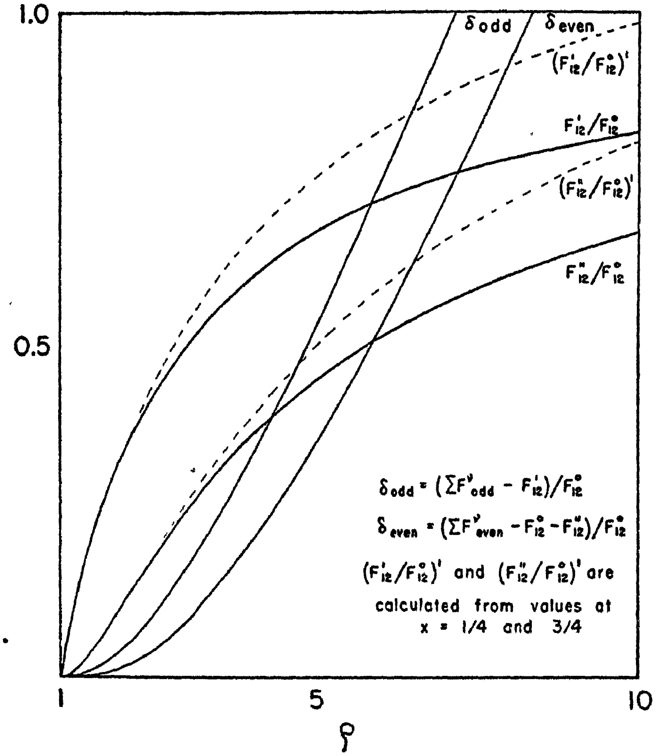


FIG. 9. Volume-symmetrical deviations

It is probable that the Flory–Huggins model represents mixtures of straight-chain hydrocarbons to a fair approximation, but most mixtures of small molecules cannot be pictured as fitting into a simple lattice. Nor can the size of one kind of molecules be neglected relative to that of another. In treating energy the lattice theory treats the number of contacts between the central molecule and those surrounding it. This should be represented approximately by the surface of the central molecule. It is not at all certain that the effective surface of a molecule is independent of the nature of the molecules around it, but it will be an improvement to make this assumption and take ρ as the ratio of surfaces instead of always the ratio of volumes. For long-chain molecules the surface is practically proportional to the volume, so ρ will still be the ratio of molal volumes. For molecules of different sizes but the same shape, the surface ratio is the two-thirds power of the volume ratio, corresponding to the formula proposed by Langmuir (13). To justify the use of a smaller ρ with the Flory–Huggins theory, we may assume that the "free volume" of a molecule is proportional to its surface, and follow Hildebrand (8) in the rest of his derivation.

This problem is of prime importance to the high-polymer and colloid chemist. Its relation to the association of small molecules is discussed later. With small molecules the place where I have found it most important is in the study of the effect of changing solvent on the rate of reaction, where the rate depends upon the ratio of the activity coefficients of the reactants to that of the critical complex.

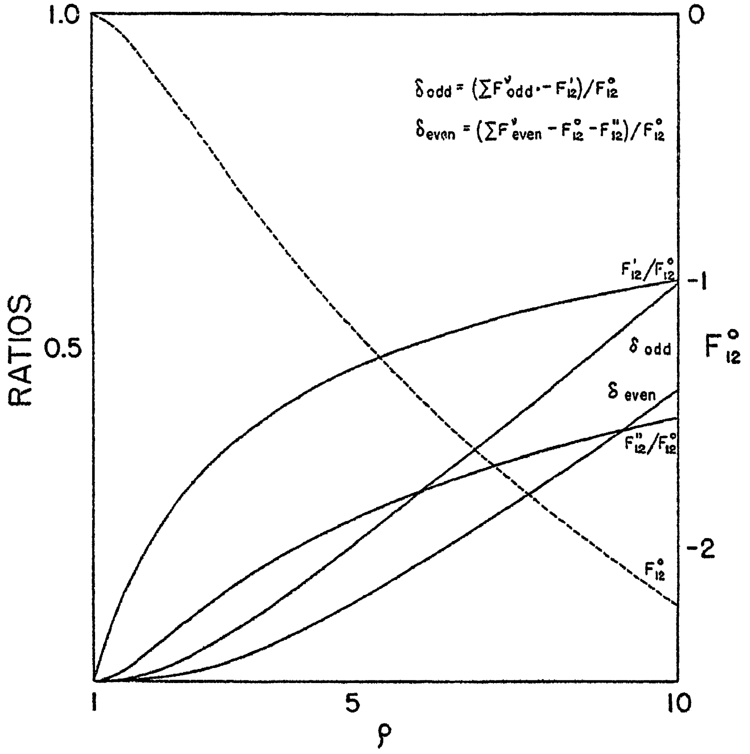


FIG. 10. Flory–Huggins deviations

The Flory–Huggins contribution is

$$\left(\ln \frac{f_A f_B}{f_x}\right)_{FH} = \ln \frac{\bar{V}_A \bar{V}_B}{\bar{V}_C V_x} + 1 - \frac{\bar{V}_A + \bar{V}_B - \bar{V}_C}{V_x}$$

If the volume of the critical complex is the sum of the volumes of the components, this becomes $-\ln V_x +$ const., which means that the rates should be compared in volume concentrations rather than in mole fractions. I have long insisted that the reverse is true, and experimental results have supported my thesis (22). If the volumes should be replaced by surfaces, the surface of the critical complex should be expected to be less than that of the reactants. If the difference is just the mean surface of a solvent molecule, the change with surface of solvent molecules is zero. It is possible that this relation is satisfied well enough so that the mole fraction expression holds approximately. It does seem a very complicated justification of a simple relation, and I am not happy about it.

EFFECTS OF VOLUME CHANGE ON MIXING

If distinction is to be made between the enthalpy and the minus temperature entropy product, it is important to consider the effect on these properties of the change of volume on mixing (19). This effect nearly cancels out in the free energy. The change in state considered in the foregoing theories, as in practically all other theories, is not the same as that treated experimentally, and it is important to consider the difference.

We shall limit ourselves to an isothermal change in state, $T = T_0$ throughout. We shall let N_A be the number of moles of component A initially at a pressure P_{A0}, with molal volume energy, entropy, enthalpy, work content, and free energy V_{A0}, E_{A0}, S_{A0}, H_{A0}, A_{A0}, and F_{A0}. The general isothermal change in state is

$$\underbrace{N_A A(P_{A0}, T_0) + N_B B(P_{B0}, T_0) + \cdots}_{\text{unmixed}} \rightarrow \underbrace{N_A A(P, T_0) + N_B B(P, T_0) + \cdots}_{\text{mixed}}$$

In the initial state

$$V_0 = N_A V_{A0} + N_B V_{B0} + \cdots\cdot ,$$

$$E_0 = N_A E_{A0} + N_B E_{B0} + \cdots\cdot , \text{ etc.}$$

In the final state

$$V = V_0 + V^M$$

$$E = E_0 + E^M, \text{ etc.}$$

There are two special cases which are so important that we shall give them special subscripts. In both the initial pressures are all the same:

$$P_{A0} = B_{B0} \cdots\cdot = P_0$$

In the constant-pressure case

$$P = P_{A0} \qquad V = V_0 + V_p^M \qquad E = E_0 = E_p^M, \text{etc.}$$

In the constant total volume case

$$V = V_0 \text{ or } V_v^M = 0 \qquad E = E_0 + E_v^M, \text{etc.}$$

Most measurements with liquids are at constant pressure. Although in measurements of vapor–liquid equilibrium, P_{A0} is the vapor pressure of pure component A, P_{B0} is that of pure component B, and P is the vapor pressure of the mixture, the correction to constant pressure is easily made and is usually much smaller than the correction due to deviations from the gas laws.

All theoretical calculations are at constant total volume. The lattice treatment assumes tacitly that the initial pressures are such that a unit of each component occupies the same volume, which may require that P_{A0} is very different from P_{B0}. More general theories will permit that they are the same, or,

$$P_{A0} = P_{B0} = \cdots = P_0$$

In this case we can calculate some very interesting relations from thermodynamics alone.

$$P - P_0 = -\int_{V_0+V^M}^{V_0} \frac{dv}{\beta v} = \frac{1}{\beta} \ln (1 + V^M/V_0) \quad \text{since } dp = -\frac{dv}{\beta v}$$

$$F_v^M - A_v^M = (P - P_0)V_0 = \frac{V_0}{\beta} \ln (1 + V^M/V_0)$$

$$F_p^M - F_v^M = \int_{V_0}^{V_0+V^M} -\frac{dv}{\beta} = -\frac{V^M}{\beta}$$

$$F_p^E - A_v^E = F_p^M - A_v^M = \frac{V_0}{\beta} [\ln (1 + V^M/V_0) - V^M/V_0]$$

$$= -(V_0/2\beta)[(V^M/V_0)^2 + \cdots]$$

$$S_p^E - S_v^E = S_p^M - S_v^M = \int_{V_0}^{V_0+V^M} \left(\frac{\partial p}{\partial T}\right)_v dv$$

$$= \int_{V_0}^{V_0+V^M} \left[\left(\frac{\partial p}{\partial T}\right)_v + \frac{d(1/\beta)}{dT} \ln \frac{V_0 + V^M}{v}\right] dv$$

$$= V^M \left(\frac{\partial P_0}{\partial T}\right)_v + \frac{V_0 \, d \ln \beta}{\beta T \, d \ln T} [\ln (1 + V^M/V_0) - V^M/V_0]$$

$$H_p^M - E_v^M = V^M T \left(\frac{\partial P_0}{\partial T}\right)_v + \frac{V_0}{\beta}\left(1 + \frac{d \ln \beta}{d \ln T}\right) [\ln (1 + V^M/V_0) - V^M/V_0]$$

$$T \left(\frac{\partial P_0}{\partial T}\right)_v = T \frac{\alpha_0}{\beta_0} = \text{“Internal pressure”}$$

$$\alpha = \left(\frac{\partial \ln V}{\partial T}\right)_p, \qquad \alpha_0 = \left(\frac{\partial \ln V_0}{\partial T}\right)_p, \qquad \beta = \left(\frac{-d \ln V}{\partial p}\right)_T,$$

$$\beta_0 = \left(\frac{-d \ln V_0}{\partial p}\right)_T$$

The terms in $[\ln (1 + V^M/V_0 - V^M/V_0]$ are usually negligibly small, and there is no other term for $F_p^E - A_v^E$. The terms in $(\partial P_0/\partial T)_v$ in the entropy and enthalpy are far from negligible. For non-polar mixtures I have calculated (19) that $V_p^M = \beta_0 F_p^E$, so that

$$H_p^M - E_v^M = T\alpha_0 F_p^E$$

For most organic liquids, α_0 is about 10^{-3}, so $T\alpha_0$ varies from 0.25 to 0.5 at the temperatures which interest us most.

INTRAMOLECULAR COHESIVE ENERGY

I found eighteen years ago (17) that the cohesive energy density of a mixture of simple non-polar liquids, defined as the difference, E/V_0, between the energy per unit volume in the liquid state, E_l/V_0, and the energy of the same material in the perfect gas state at the same temperature, E_g/V_0, should be a quadratic function of the volume fractions of the components, and that the coefficient of each term is the product of the square roots of the cohesive energy densities of the two components involved in that term.

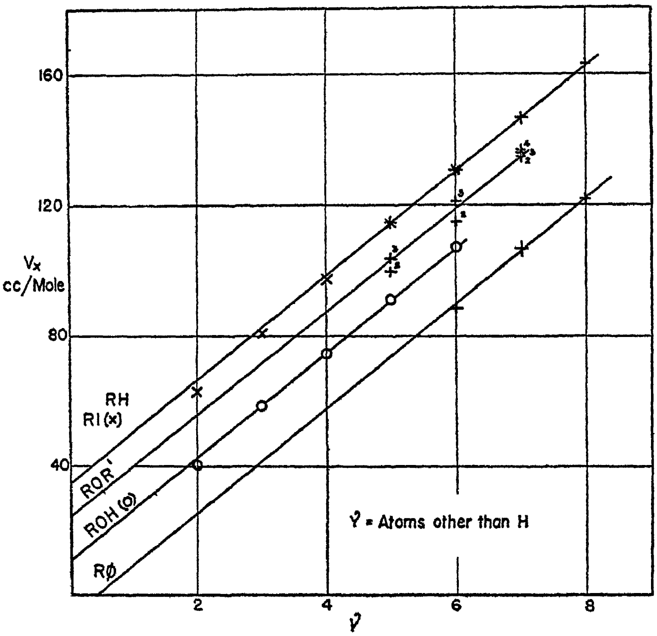


Fig. 11. Normal aliphatic compounds: volume

$$E/V_0 = (E_l - E_g)/V_0 = \Sigma_{jk}N_jV_{j0}N_kV_{k0}a_{jk}/V_0^2$$

$$= \Sigma_j N_j V_{j0}\sqrt{a_{jj}}\;\Sigma_k N_k V_{k0}\sqrt{a_{kk}}/V_0^2 = (\Sigma_j N_j V_{j0}\sqrt{a_{jj}}/V_0)^2$$

$$a_{jk} = \sqrt{a_{jj}\,a_{kk}} \qquad a_{jj} = E_j/V_{j0}$$

Last year[3] I made the suggestion that long flexible molecules should have an intramolecular cohesive energy in the perfect gas state, so that the energy available for interaction is larger than the difference between the liquid and the perfect gas.

---

[3] This section was presented at the 111th Meeting of the American Chemical Society in Atlantic City, April 17, 1947, but has not been previously published.

To study this effect let us first consider the different parts of a single molecule as components. Then members of a homologous series correspond to mixtures of different composition. Testing first the additivity of volumes, figure 11 shows the volumes of several normal aliphatic compounds plotted against ν, the number of atoms other than hydrogen. The lines are all drawn with slopes of 16 ml. per carbon atom, which appears to be the volume occupied at 20°C. by a CH_2 group in straight chains. This volume is also almost independent of the temperature, and the thermal expansion is almost entirely in the intercepts, which differ for different series. The top curve represents hydrocarbons and iodides, the second the ethers, the third alcohols and the fourth phenyl derivatives.

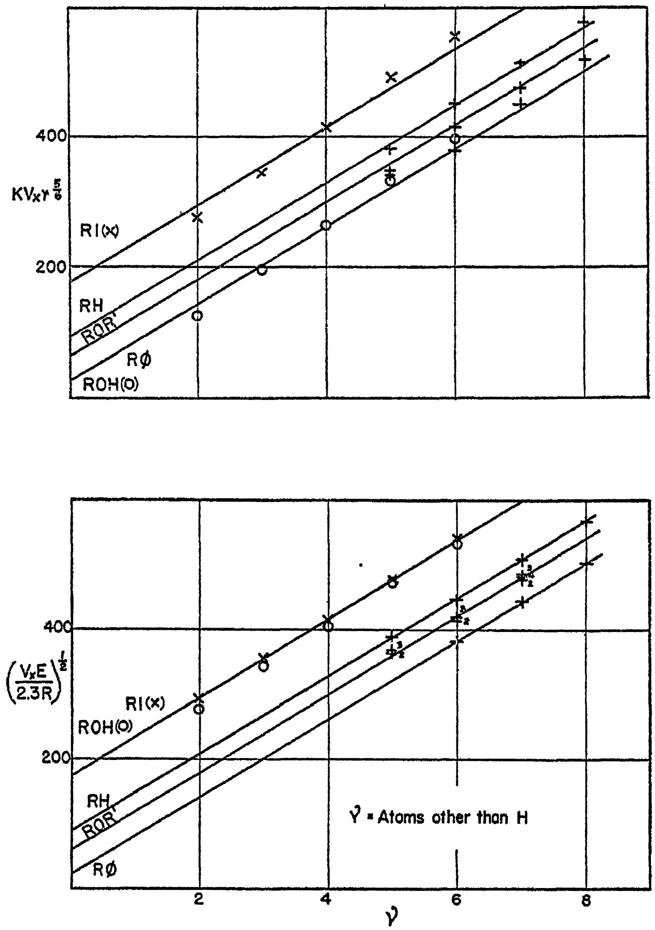


FIG. 12. Normal aliphatic compounds: (upper) refraction–volume; (lower) cohesive energy–volume.

The lower part of figure 12 shows $(V_0E/2.3)^{1/2}$ for these same substances. This should be a linear function of ν if the cohesive energy density is a quadratic function and $a_{jk} = \sqrt{a_{jj}a_{kk}}$, for

$$V_0E = V_0^2(E/V_0) = (\Sigma_j N_j V_{j0}\sqrt{a_{jj}})^2$$

The lines all have the slope of 60 units per carbon atom. The series are again characterized by the intercepts. The top curve represents the iodides, the second the hydrocarbons, the third the ethers, and the fourth the phenyl derivatives. The alcohols lie near the iodides and could be fitted by a straight line with slope slightly greater than that of the others. The slope would be increased if the interaction of the end groups, methyl and hydroxyl, with methylene groups is slightly greater than the geometric mean of the interaction of end groups with end groups and methylenes with methylenes.

The upper part of figure 12 shows the cohesive energy densities calculated from the refraction, $(n^2 - 1)/(n^2 + 2)$, for the sodium D line. The dispersion or London energy of two molecules depends upon the polarizabilities and is usually calculated as depending upon some other function of the mobility of the electrons. We shall assume that this part is also a function of the polarizability,

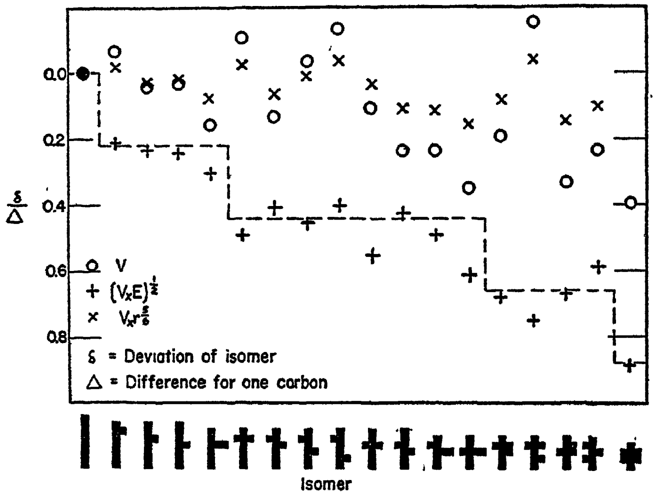


FIG. 13. Octanes: deviations from normal divided by change for one carbon

and that the cohesive energy density is proportional to some power of the refraction. Empirically I found the five-thirds power, and chose a coefficient to fit an average for the non-polar substances in this figure. The lines are taken directly from the lower figure.

We notice first that the alcohols have dropped from the top curve to the bottom. This is not surprising, since a large part of their cohesion comes from the exposed dipoles or hydrogen bonds. We notice next that every series indicates a slope slightly greater than that of the lines. This difference may be due to errors in our assumptions, but it cannot be corrected by a reasonable change of the exponent of the polarizations. If we assume that it is real, the difference indicates that the energy of evaporation gives too small a value for the cohesive energy of molecules long and flexible enough to show intramolecular cohesion in the perfect gas state.

Figure 13 shows the volume and the square root of the volume cohesive energy product calculated by the two methods for the isomeric octanes. The formulas of the isomers are given at the bottom, going from *n*-octane through 2-, 3-, and 4-methylheptanes, etc., to 2,2,3,3-tetramethylbutane. Except for this last substance the measurements are from Dr. Rossini's laboratory in the United States Bureau of Standards (27). Each function is represented by the difference from *n*-octane divided by the difference of *n*-nonane from *n*-octane.

The circles represent the volumes. I find little regularity except that a branch on the 2 carbon tends to increase the volume, and other branches tend to decrease it, particularly two branches on adjacent carbons. One might expect more regularity in these volumes. They are measured at atmospheric pressure, but an increase of only 0.01 per cent gives the volume under no external field with negligible differences between isomers.

The root energy–volume products calculated from refractions follow the volumes to a certain extent, but show a smaller deviation. The measured root energy–volume products show a very different behavior. The differences are greater and also more regular. The broken line in the figure drops an equal amount for each branch in the chain regardless of its position. It represents the experimental points to a fair approximation.

The evidence is clear that there is more intramolecular cohesive energy in the perfect gas state for branched chains, although it is not proven that this energy is available for intermolecular cohesion in the liquid. If it is available, the refractions should give a better measure of the behavior of solutions than the cohesive energies themselves. At the worst, the refractions should serve as an approximate measure of the cohesive energy density when direct measurements are impossible because of high melting point or decomposition on vaporization. For such substances the refraction, and if necessary the volume, may be measured in solution.

SORTING OF MOLECULES

Of course the distribution of the molecules in a solution will not be random if there is a change of energy when contact between two A molecules and that between two B molecules is replaced by two contacts between A and B molecules. The effect of this deviation on the thermodynamic properties has been calculated by Rushbrooke (16), by Kirkwood (12), and by Guggenheim (7) for spheres of equal volume with z nearest neighbors. Kirkwood's expression is

$$E^M_{vx} = B_{12}x_1x_2 - (2B^2_{12}/zRT)x^2_1x^2_2$$

with B_{12} independent of the temperature:

$$S^E_{vx} = -(B^2_{12}/zR)x^2_1x^2_2$$

If there were no other cause for deviation from the simple theory, we would have:

$$F^E_{xp} = B_{12}x_1x_2[1 - (B_{12}/zRT)x_1x_2]$$

In terms of the expansions used above we may say, since $x_1x_2 = \dfrac{1 - (x_1 - x_2)^2}{4}$, $E^0_{v12} = B_{12}(1 - B_{12}/2zRT)$, $E''_{v12} = B^2_{12}/2zRT$, $S^0_{v12} = -B^0_{12}/4zRT$, $S''_{v12} =$

$B_{12}^2/4zRT$, $F^0_{p12} = B_{12}(1 - B_{12}/4zRT)$, $F''_{p12} = B_{12}^2/4zRT$. The deviation is always negative. If B_{12} is positive and z is 8, the deviation is small until there is separation into two liquid phases when

$$\frac{B_{12}}{RT}\left(1 - \frac{B_{12}}{2zRT}\right) = 2$$

This treatment is based on the assumption that the energy of interaction is independent of the orientation. The theory of polar molecules would be advanced greatly if we had a similar treatment in which the interactions vary from spot to spot on the molecules. In the meantime we must be satisfied with treatments in terms of combination or association which do not distinguish between energy and enthalpy. By combination we mean interaction between unlike molecules which can be expressed by the law of mass action; by association we mean similar interaction between like molecules. Three types of association are important. In the first there is only one spot of special reactivity in each molecule, so the association is limited to dimers. We shall call this "acid-type association." In the second type, which we shall call "alcohol-type association," two active spots on each molecule permit the formation of linear chains of any length. In the third type, which we shall call "water-type association," four active spots on each molecule permit the formation of three-dimensional polymers. It is convenient to discuss the general case in which the number of reactive spots in each molecule is two or more. We shall discuss only the simplest type of combination, in which each molecule has but one spot which can react with a spot in a molecule of the other species.

We shall not make the assumption that chemical action is the only cause for deviations from the ideal solution laws. It has been known for many years (26) that no combination of reactions which obey the law of mass action can lead to separation into two liquid phases. We shall follow Redlich and Kister (14) in assuming that the non-chemical deviations do not depend upon the complexity of the molecules. It follows that the equilibrium constants do not depend upon the composition of the system.

We shall use the subscript A for the component which associates and B for the other; n_A and n_B are the numbers of formula weights of the components; x_A and x_B are the stoichiometric mole fractions, $x_A = n_A/(n_A + n_B)$; ν_g is the number of moles of the species g; and ξ_g and ξ_B are the "true" mole fractions, $\xi_g = \nu_g/(n_B + \Sigma_g \nu_g)$. The parts of the chemical potentials due to association, or combination, are

$$\mu_A^a/RT = \ln \gamma_A/\beta_A = \ln \xi_1/x_A\xi_1^0$$

$$\mu_B^a/RT = \ln \gamma_B/\beta_B = \ln \xi_B/x_B = \ln \xi_B/(1 - x_A)$$

in which γ_A and γ_B are the stoichiometric activity coefficients, β_A and β_B are those parts which are independent of the amount of chemical action, and ξ_1^0 is the value of ξ_1 when $x_A = 1$. The free energy per mole of components due to chemical action is

$$F_x^a/RT = x_A \ln \xi_1/x_A\xi_1^0 + (1 - x_A) \ln \xi_B/(1 - x_A)$$

It is convenient to work with one mole of total species, and very convenient to take ξ_1 as the independent variable.[4] Then it is necessary to determine ξ_1^0 as a function of K, and x_A and ξ_B as functions of K and ξ_1. We shall also use $Q_x = F_x^E/2.3RT$, etc. and common logarithms.

ACID-TYPE ASSOCIATION

In the acid-type association there are but three species, B and the monomers and dimers of A, and

$$\xi_2 = K\xi_1^2$$

$$n_A = \xi_1 + 2\xi_2 = \xi_1(1 + 2K\xi_1)$$

$$n_B = \xi_B = 1 - \xi_1(1 + K\xi_1)$$

$$x_A = \xi_1(1 + 2K\xi_1)/(1 + K\xi_1^2)$$

$$\xi_1^0 = (\sqrt{1 + 4K} - 1)/2K$$

So

$$Q_x^a = x_A \log \frac{(1 + K\xi_1^2)2K}{(1 + 2K\xi_1)(\sqrt{1 + 4K} - 1)} + (1 - x_A) \log (1 + K\xi_1^2)$$

The expression for x_A is a quadratic and it is possible to solve for ξ_1 and to express Q_x explicitly in x_A, but the expression is probably too complicated for general use.

Since

$$(\bar{Q}_A^a)_B = -\log \xi_1^0$$

and

$$(\bar{Q}_B^a)_A = \log (2 - \xi_1^0)$$

it is easy to obtain the results

$$\Sigma Q_{AB\ \mathrm{even}} = \tfrac{1}{2} \log (2/\xi_1^0 - 1) = (1 - \xi_1^0)/2.3 + \cdots$$

and

$$\Sigma Q_{AB\ \mathrm{odd}} = \tfrac{1}{2} \log \xi_1^0(2 - \xi_1^0) = (1 - \xi_1^0)^2/2.3 + \cdots$$

When $x_A = \frac{1}{2}$

$$\xi_1 = \sqrt{1 + 3K} - 1)/3K$$

So

$$Q_{12}^0 = 2 \log \frac{4}{27} \frac{(1 + 6K - \sqrt{1 + 3K})^2}{K(2\sqrt{1 + 3K} + 1)(\sqrt{1 + 4K} - 1)}$$

EXTENDED ASSOCIATION

For the alcohol type Redlich and Kister (14) follow previous workers in assuming that the association constant, K, for the reaction

$$A + A_{n-1} = A_n$$

---

[4] This is equivalent to taking temperature, pressure, and the potential (or activity) of component 1 as independent variables and closer to Gibbs than our usual procedures.

does not depend upon the order of the association. They obtain results which may be summed as

$$\bar{Q}^{\alpha}_{\mathrm{A}} = \log\,[(1 - K\xi_1)^2 + K\xi_1^2](1 + K)$$

$$\bar{Q}^{\alpha}_{\mathrm{B}} = \log\,[(1 - K\xi_1)^2 + K\xi_1^2]/(1 - K\xi_1^2)$$

$$Q^0_{\mathrm{AB}} = 4\log\left[1 + \frac{(\sqrt{1+K} - 1)^2}{K}\right] = 4\log\frac{2(\sqrt{1+K} - 1)}{K\sqrt{1+K}}$$

$$\Sigma Q_{\mathrm{AB\ even}} = \log\,(1 + K)$$

$$\Sigma Q_{\mathrm{AB\ odd}} = 0$$

The symmetry of their results depends upon the expression of the equilibria in terms of mole fractions. The association polymers must give deviations due to different sizes of the different polymers. If we assume that the potentials are given by the Flory–Huggins expressions, the constants should be expressed in terms of volume fractions or concentrations (4).

We follow, in so far as we need to go, Stockmayer's method (24) of treating branched polymers. We consider a system of N monomer units of component A, existing as M molecules, and N_{B} molecules of component B which does not associate. We let f be the number of reactive spots on each unit, and α be the fraction of those spots which have reacted. From the assumption that there are no rings, $M = N(1 - \alpha f/2)$, and all of component A is in a single molecule when $\alpha f/2 = 1 - \frac{1}{N}$. From the assumption that every reactive group has the same probability of reacting, the number of monomer and dimer units, m_1 and m_2 are given by

$$m_1/N = (1 - \alpha)^f$$

$$m_2/N = f\alpha(1 - \alpha)^{2f-2}/2$$

Stockmayer finds that his expression for the number of a large polymer becomes invalid when $\alpha f/2 = f/2(f - 1)$, but there seems to be no reason why those parts which we have taken should not hold to the limit $\alpha f/2 = 1$.

We use the equilibrium expression for the formation of the dimer

$$\frac{m_2}{m_1^2} = \frac{f\alpha}{2N(1 - \alpha)^2}$$

But by the Flory–Huggins expression

$$\frac{m_2}{m_1^2} = \frac{k_2}{N + \rho N_{\mathrm{B}}} \quad \text{if } \rho = V_{\mathrm{B}}/V_1$$

To find k_2 we go to the limit $\alpha = \alpha_0$ when $N_B = 0$

$$k_2 \quad \frac{f}{2}\,\frac{\alpha_0}{(1-\alpha_0)^2}$$

$$\frac{N}{N + \rho N_B} = \frac{x_A}{1 + (\rho - 1)x_B} = \frac{\alpha(1-\alpha_0)^2}{(1-\alpha)^2\alpha_0}$$

It is convenient to use α_0 as the measure of the tendency to associate and to use α as the working variable. With the Flory–Huggins expression for the activity coefficients, we obtain

$$\bar{Q}^a_A = \log\left(\frac{1-\alpha}{1-\alpha_0}\right)^f - \log\left[(1 + x_B(\rho - 1)\right] + \frac{1}{2.3}\left[1 - \alpha_0 f/2 - \frac{1 - x_A\alpha f/2}{1 + x_B(\rho - 1)}\right]$$

$$\bar{Q}^a_B = \log\rho - \log\left[1 + x_B(\rho - 1)\right] + \frac{1}{2.3}\left[1 - \rho\frac{1 - x_A\alpha f/2}{1 + x_B(\rho - 1)}\right]$$

If $\rho = 1$, $\qquad x_A = \dfrac{\alpha(1-\alpha_0)^2}{\alpha_0(1-\alpha)^2}$

$$\bar{Q}^a_A = \frac{f}{2}\left[\log\left(\frac{1-\alpha}{1-\alpha_0}\right)^2 + \frac{1}{2.3}(x_A\alpha - \alpha_0)\right]$$

$$\bar{Q}^a_B = \frac{f}{2}\frac{1}{2.3}x_A\alpha$$

$$Q^a_0 = 4\frac{f}{2}\left[\log\frac{1-\alpha_{0.5}}{1-\alpha_0} + \frac{1}{4.6}(\alpha_{0.5} - \alpha_0)\right]$$

$$\Sigma Q^a_{AB\,\text{even}} = -\frac{f}{2}\log(1-\alpha_0)$$

$$\Sigma Q^a_{AB\,\text{odd}} = \frac{f}{2}\left[\log(1-\alpha_0) + \frac{1}{2.3}\alpha_0\right]$$

If we assume the mole fraction form, on the other hand,

$$K = \frac{m_2}{m_1^2}(M + N_B) = \frac{f}{2}\frac{\alpha_0(1-\alpha_0 f/2)}{(1-\alpha_0)^2}$$

For $f = 2$, $\qquad K = \dfrac{\alpha_0}{1-\alpha_0}$

$$Q^0_{AB} = 4\log 2(1 - \sqrt{1-\alpha_0})$$

$$\Sigma Q_{AB\,\text{even}} = -\log(1-\alpha_0)$$

$$\Sigma Q_{AB\,\text{odd}} = 0$$

To show these relations graphically figure 14 gives $Q^0_{AB}2/f$, $-\ \Sigma Q_{AB\,\text{odd}}2/f$ and $(\Sigma Q_{AB\,\text{even}} - Q^0_{AB})2/f$ plotted against α_0 for the Flory–Huggins expression of extended association with $\rho = 1$. The broken lines show the corresponding

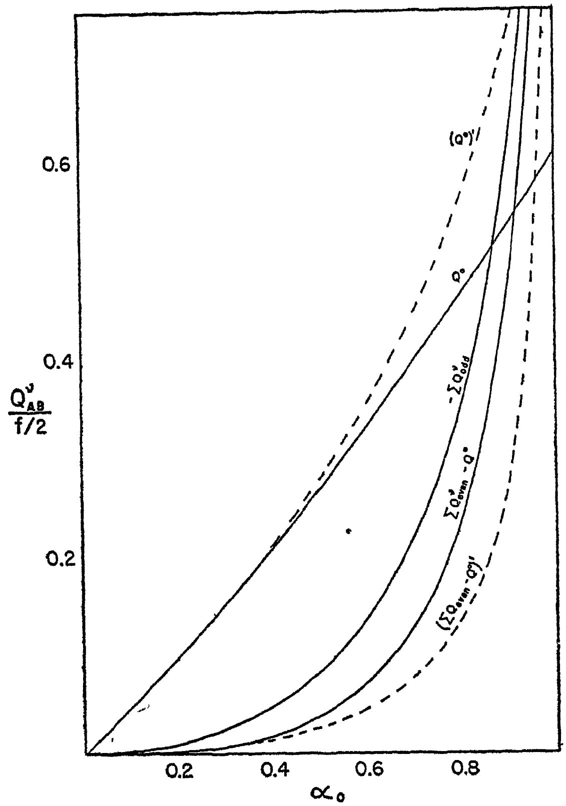


FIG. 14. Extended association

functions for the mole fraction expression for $f = 2$. In this case each Q_{odd} is zero.

Figure 15 shows the Q^0_{AB}, $-\ \Sigma Q_{odd}$, and $(\Sigma Q_{even} - Q^0_{AB})$ as functions of α_0 for the acid type of association in terms of mole fractions. The difference from figure 14 is not very great.

In figure 14 the curves for each value of f stop just before $\alpha_0 = 2/f$. Since $Q_0 2/f$ is approximately proportional to α_0, the value of Q_0 is about the same for all values of f at the same value of $\alpha_0 f/2$, that is, for an equal number of linkages per unit or an equal degree of association. The higher terms become relatively less important as f increases.

It is most interesting to note that the disappearance of the odd terms for the alcohol-type association depends upon the use of the mole fraction form, as well as upon the nature of the association. The quantitative values given in figures 14 and 15 are probably inaccurate because of the errors in the Flory–Huggins equations, but there must be enough truth in their picture to show that this type of association should not lead to symmetrical expressions. It must be remembered, moreover, that both treatments depend on the assumption that there is no ring formation. The probability of this assumption decreases as f increases, but it must be in error even for $f = 2$.

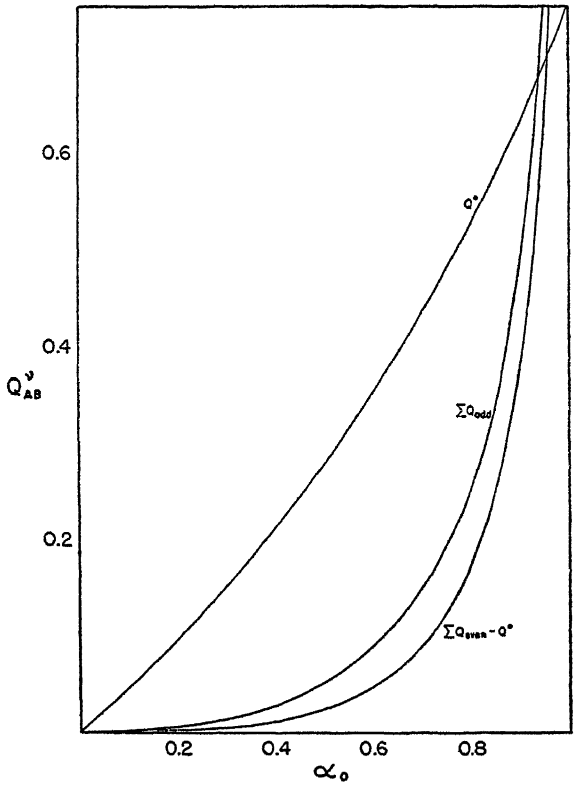


FIG. 15. Acid-type association

ONE-TO-ONE COMBINATION

If we have the reaction

$$A + B = C$$

$$\xi_C = K\xi_1\xi_B = 1 - \xi_1 - \xi_B = \frac{K\xi_1(1 - \xi_1)}{1 + K\xi_1}$$

$$\xi_B = \frac{1 - \xi_1}{1 + K\xi_1}$$

$$n_A = \xi_1 + \xi_C = \frac{\xi_1(1 + K)}{1 + K\xi_1}$$

$$n_B = \xi_B + \xi_C = 1 - \xi_1$$

$$x_A = \frac{\xi_1(1 + K)}{1 + 2K\xi_1 - K\xi_1^2}$$

$$Q^c_x = x_A \log \frac{1 + K\xi_1 - K\xi_1^2}{1 + K} + (1 - x_A) \log \frac{1 + 2K\xi_1 - K\xi_1^2}{(1 + K\xi_1)^2}$$

$$(\bar{Q}^c_A)_B = -\log (1 + K)$$

$$(\bar{Q}^c_B)_A = -\log (1 + K)$$

$$\Sigma Q^{\nu}_{AB\ \text{even}} = -\log (1 + K) = -2 \log (1 + \alpha_{0.5})(1 - \alpha_{0.5})$$

in which $\alpha_{0.5}$ is the fraction of groups reacted when $x = 0.5$. In the case of one-to-one combination, the deviations must obviously be symmetrical, so each Q_{odd} is zero. When $x_A = \frac{1}{2}$,

$$\xi_1 = \xi_B = (\sqrt{1 + K} - 1)/K$$

So

$$Q^0_{AB} = 4 \log 2(\sqrt{1 + K} - 1)/K = 4 \log (1 - \alpha_{0.5})$$

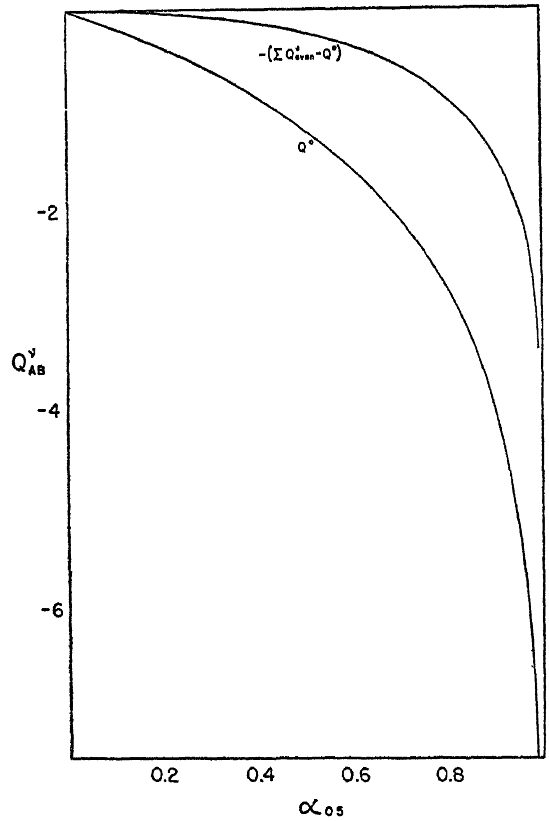


FIG. 16. One-to-one combination

Figure 16 shows Q^0_{AB} and $-(\Sigma Q_{AB\ \text{even}} - Q^0_{AB})$ as a function of $\alpha_{0.5}$. The scale of the ordinates is ten times as large as in figures 14 and 15. Since the higher terms have the opposite sign to Q_0, the curve is more pointed than the parabola of a symmetrical solution.

If each type of molecule has a single reactive group, we should expect dimers of each component and the one-to-one compound. If either has more than one group, more complicated behavior may be expected. The free energies illustrated in figures 7 and 8 may be approximated by assuming the formation of two complexes $(C_2H_5OH)_2$ and $CHCl_3(C_2H_5OH)_3$, and doubtless could be approximated by linear chains of alcohol molecules branched, but with weaker links, at chloroform molecules. It is probable that the behavior of any completely miscible solutions can be expressed by chemical reactions obeying the law of mass action. However, it is certain that no partial miscibility can be explained in this way. Moreover, the assumption that the equilibrium constants are independent of the medium is a most improbable one, for the reactive groups such as hydroxyl must also interact with non-polar groups, and differently with different groups. Finally, all the other causes of deviations from ideal laws discussed earlier must also operate in systems in which there are chemical reactions. The best advice which comes from years of study of liquid mixtures is to use any model in so far as it helps, but not to believe that any moderately simple model corresponds very closely to any real mixture.

REFERENCES

(1) Benedict, M., Johnson, C. A., Solomon, E., and Rubin, L. C.: Trans. Am. Inst. Chem. Engrs. **41,** 371 (1945).
(2) Carlson, H. C., and Colburn, A. P.: Ind. Eng. Chem. **34,** 581 (1942).
(3) Flory, P. J.: J. Chem. Phys. **9,** 660 (1941); **10,** 51 (1942); **13,** 453 (1945).
(4) Flory, P. J.: J. Chem. Phys. **12,** 425 (1944).
(5) Gibbs, J. W.: Trans. Conn. Acad. Arts Sci. **3,** 115 (1875); *Collected Works*, Vol. I, p. 62.
(6) Guggenheim, E. A.: Trans. Faraday Soc. **33,** 151 (1937).
(7) Guggenheim, E. A.: Proc. Roy. Soc. (London) **A183,** 203, 213 (1944); Trans. Faraday Soc. **41,** 107 (1945).
(8) Hildebrand, J. H.: J. Am. Chem. Soc. **51,** 69 (1929).
(9) Hildebrand, J. H.: J. Chem. Phys. **15,** 225 (1947).
(10) Hildebrand, J. H.: Chem. Revs. **44,** 37 (1949).
(11) Huggins, M. L.: J. Chem. Phys. **9,** 440 (1941): J. Phys. Chem. **46,** 151 (1942); Ann. N. Y. Acad. Sci. **41,** 11 (1942); J. Am. Chem. Soc. **64,** 1712 (1942).
(12) Kirkwood, J. G.: J. Phys. Chem. **43,** 97 (1939).
(13) Langmuir, I.: Third Colloid Symposium Monograph, p. 3. The Chemical Catalog Company, Inc., New York (1925).
(14) Redlich, O., and Kister, A. T.: J. Chem. Phys. **15,** 849 (1947).
(15) Redlich, O., and Kister, A. T.: Ind. Eng. Chem. **40,** 345 (1948).
(16) Rushbrooke, G. S.: Proc. Roy. Soc. (London) **A166,** 296 (1938).
(17) Scatchard, G.: Chem. Revs. **8,** 321 (1931).
(18) Scatchard, G., and Hamer, W. J.: J. Am. Chem. Soc. **57,** 1805 (1935).
(19) Scatchard, G.: Trans. Faraday Soc. **33,** 160 (1937).
(20) Scatchard, G., and Raymond, C. L.: J. Am. Chem. Soc. **60,** 1278 (1938).
(21) Scatchard, G., Wood, S. E., and Mochel, J. M.: J. Phys. Chem. **43,** 119 (1939).
(22) Scatchard, G.: J. Chem. Phys. **7,** 657 (1939).
(23) Scatchard, G., Wood, S. E., and Mochel, J. M.: J. Am. Chem. Soc. **62,** 712 (1940).
(24) Stockmayer, W. H.: J. Chem. Phys. **11,** 45 (1943).
(25) van Laar, J. J.: Z. physik. Chem. **72,** 723 (1910); **83,** 599 (1913).
(26) Washburn, E. W.: Trans. Am. Electrochem. Soc. **22,** 330 (1912).
(27) Willingham, C. B., Taylor, W. J., Pignocco, J. M., and Rossini, F. D.: J. Research Natl. Bur. Standards **35,** 219 (1945).
(28) Zimm, B. H.: J. Chem. Phys. **14,** 164 (1946).

A CRITIQUE OF THE THEORY OF SOLUBILITY OF NON-ELECTROLYTES[1]

JOEL H. HILDEBRAND

Department of Chemistry, University of California, Berkeley, California

Received October 25, 1948

The present state of theory is indicated by a classification of solution types based upon deviations from ideal (zero) heat of mixing, ideal entropy of mixing due to unequal molal volumes, and excess attraction between like molecules (association) or unlike molecules (solvation). The quantitative formulation of the effects of unequal molecular attractions and sizes has been made by the aid of simplified models, and sources of error and uncertainty include: (*1*) the assumption that intermolecular forces are radial from molecular centers, (*2*) the neglect of dipole interactions, which are especially large in cases of hydrogen bonding, (*3*) oversimplification of repulsive forces, (*4*) the Berthelot relation between attraction constants, (*5*) the assumption that thermal agitation is sufficient to give completely random distribution despite unequal attractive forces, (*6*) neglect of volume changes on mixing liquids, (*7*) uncertainty regarding the values of heat of vaporization at various temperatures, (*8*) in the case of solid solutes, the extrapolation of liquid properties below the melting point, (*9*) uncertainty regarding the entropy of mixing molecules of different sizes and shapes, and (*10*) chemical interactions.

The magnitude of the net uncertainty attending solubility calculations for regular solutions is appraised by calculating solubility parameters from actual measurements for a number of interrelated systems and noting their mutual consistence.

The present state of the theory of solubility of non-electrolytes may perhaps be best appreciated by the aid of the scheme in table 1. This represents a classification of the main essential factors rather than of solutions themselves, because scarcely any actual solution can be said to behave solely in any one of the several ways designated by such terms as "ideal" or "regular."

Solubilities express the composition of phases in equilibrium, and hence are to be calculated as free energies. Recent attacks upon the theoretical problem have split it into its two natural components, heat and entropy. The "ideal" or "perfect" solution is formed from its pure component liquids with zero heat of mixing at all temperatures and its components must accordingly have a partial molal entropy of transfer from pure liquid to solution of $\bar{s}_2 - s_2^0 = -R \ln x_2$, for component 2 and similarly for component 1. Such a solution obeys Raoult's law, which is that the activity of a component equals its mole fraction, $a_2 = x_2$. It has only recently been fully recognized that the components must have substantially equal molal volumes in order for Raoult's law to hold strictly, except in the special case of molecules such as two normal paraffins of different lengths in parallel array (3, 10). However, there can be considerable difference between the molal volumes of the pure components without causing serious departures from Raoult's law, because the radii of two molecules will be much more nearly equal than their volumes and therefore a considerable volume discrepancy can exist without greatly affecting the structure of the solution.

---

[1] Presented at the Symposium on Thermodynamics and Molecular Structure of Solutions, which was held under the auspices of the Division of Physical and Inorganic Chemistry at the 114th Meeting of the American Chemical Society, Portland, Oregon, September 13 and 14, 1948.

The possibility of designating many non-ideal solutions as "regular" was pointed out by me many years ago (2). These are solutions in which thermal agitation is sufficient to overcome the segregating effect of unequal molecular attraction and give a structure with maximum disorder or randomness, and a partial molal entropy considered originally to be ideal, i.e., $-R \ln x_2$. If the structure can be assumed to be thus random, it is possible to integrate the intermolecular potentials of all the like and unlike molecular pairs over the whole solution and, invoking the Berthelot relation, $a_{12} = \sqrt{a_{11}a_{22}}$, for the constants of attraction between the two species, to express the heat of mixing in terms of the molal volumes and energies of vaporization of the pure components. The expression for the partial molal heat of mixing of component 2 is given in table 1, where φ_1 denotes the volume fraction of component 1 and δ_1 and δ_2 are what I shall call the "solubility parameters" of the pure liquids; they are the square roots of their energies of vaporization per cubic centimeter, i.e., $\delta_1 = (\Delta E_1^V/V_1)^{1/2}$, where ΔE^V and V are molal quantities. The solubility equation

TABLE 1

Classification of solutions upon basis of enthalpy and entropy involved in transferring 1 mole of component 2 from its pure liquid to a solution in which its mole fraction is x_2

DESIGNATION	$\bar{H}_2 - H_2^0$	$\bar{S}_2 - S_2^0$	REMARKS
1. Athermal, ideal..	0	$-R \ln x_2$	$a_2 = x_2$, Raoult's law, $V_1 = V_2$
2. Regular......	$V_2\phi_1(\delta_1 - \delta_2)^2$	$-R \ln x_2$	$a_2 > x_2$, $V_1 \cong V_2$, $a_{12} = \sqrt{a_{11}a_{22}}$
3. Athermal, non-ideal...	0	$-R\left[\ln \phi_2 + \phi_1\left(1 - \frac{V_2}{V_1}\right)\right]$	$a_2 < x_2$, $V_1 \neq V_2$, $a_{12} = \sqrt{a_{11}a_{22}}$
4. General, no specific interaction...	0	$-R\left[\ln \phi_2 + \phi_1\left(1 - \frac{V_2}{V_1}\right)\right]$	$a_{12} = \sqrt{a_{11}a_{22}}$
5. One component associated... ..	Positive	$> (-R \ln x_2)$	$a_{12} < \sqrt{a_{11}a_{22}}$
6. Solvated.....	Negative	$< (-R \ln x_2)$	$a_{12} > \sqrt{a_{11}a_{22}}$

for a regular solution is obtained by combining the heat and entropy terms to give[2]

$$RT \ln (a_2/x_2) = v_2 \varphi_1^2(\delta_1 - \delta_2)^2 \tag{1}$$

I shall return to this equation after discussing briefly the other types referred to in table 1.

The question of the entropy of mixing molecules of different sizes and shapes has recently received much attention, stimulated immensely by the developments in the field of high polymers. It has led to the equation derived independently by Flory and Huggins according to which the partial molal entropy of mixing is (5)[3]

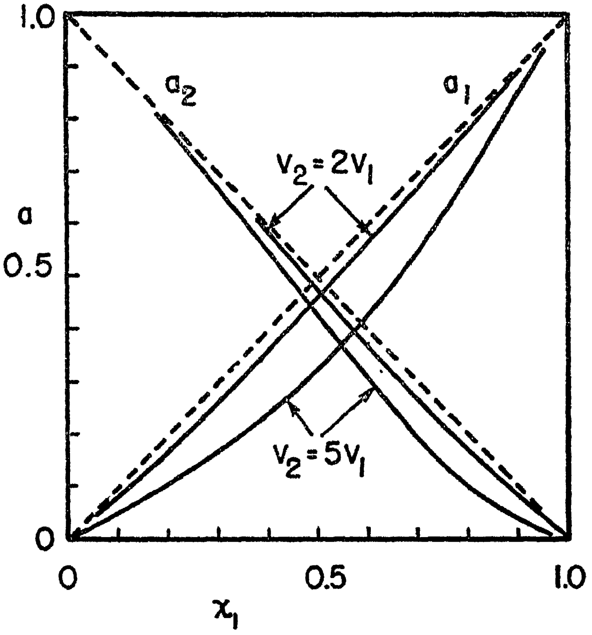


FIG. 1. Effect of inequality in molal volumes of the components of a binary solution upon their activities.

$$\bar{s}_2 - s_2^0 = -R\left[\ln \phi_2 + \phi_1\left(1 - \frac{v_2}{v_1}\right)\right] \tag{2}$$

The difference between this and $R \ln x_2$ is small for volume ratios of the order of 2:1, which are seldom exceeded by non-polar liquids with which we ordinarily deal. For example v = 61 cc. for carbon disulfide and 132 cc. for *n*-hexane at 25°C. The sign and magnitude of the deviations from Raoult's law given by equation 2 for volume ratios of 2 and 5 are illustrated in figure 1. There is some reason to believe that the equation overdoes the correction. It is to be noted that whereas the positive heat of mixing of regular solutions corresponds to positive deviations from Raoult's law, i.e., $a_2 > x_2$, the effect of inequality in molal volumes is to produce negative deviations, i.e., $a_2 < x_2$.

---

[2] For an account of the development of this equation and its underlying theory see reference 9.

[3] For a résumé of the development of this theory see Hildebrand (5).

Since most non-polar solutions in which specific or chemical interactions are absent would present both volume discrepancy and positive heat of mixing, one may combine equations 1 and 2 to give, for class 4 in table 1,

$$RT \ln (a_2/\varphi_2) = \mathrm{v}_2\varphi_1^2(\delta_1 - \delta_2)^2 + RT\,\varphi_1(1 - \mathrm{v}_2/\mathrm{v}_1) \tag{3}$$

If one component is associated, as by hydrogen bonding, its dilution by the other causes heat absorption and we have chemical equilibria to deal with. The effect is to give positive deviations from Raoult's law (class 5). An enhanced attraction between the unlike species, on the other hand, is accompanied by positive heat, diminished entropy, and negative deviations from Raoult's law (class 6).

Actual solutions may, of course, present combinations of the above effects which may be difficult or even impossible to disentangle. The present critique is aimed at an appraisal of the adequacy of equations 1 and 3 to solutions where there is reason to believe that the highly specific chemical effects referred to in classes 5 and 6 are practically absent. Moreover, since equation 2 has not been thoroughly tested, we shall concern ourselves mainly with components for which the molal volume ratios do not exceed 2 and apply equation 1.

Let us begin by reviewing the various sources of error or uncertainty. In deriving the above equation, the following assumptions were either stated or at least implicit. I say implicit, because those of us who had a hand in it were not always conscious at the time of their precise import. (*1*) The intermolecular forces operating are the short-range additive "dispersion forces," for the understanding of which we are indebted to London. If the molecules possess dipoles, strict additivity can no longer be assumed. London (12), in an important paper on "Centers of van der Waals Attraction," has discussed two sources of non-radial behavior, stating that "(1) The elementary units of the dipole interactions are, in general, not at all spherically symmetrical forces. They have rather to be built up by highly *anisotropic force centers.* (2) In certain large molecules we encounter characteristic, long, extended electronic oscillators. In these molecules the *spacial extension* of the oscillators has to be accounted for, ... and it is suggested in these cases that the molecular forces be built up by certain smaller units, which forces, however, are *no longer additive.*"

(*2*) These force fields are considered as radial and of similar form. This is an obvious oversimplification for molecules such as paraffins, geometrically far from being spherical. Furthermore, Gilman and I (7) have presented evidence that two molecular species such as ethane and 1,2-dimethylbutane, which are closely similar except for considerable difference in size, do not conform to the theory of corresponding states, as they should if their force fields were like those of the monatomic gases.

The errors introduced by neglecting dipoles depend upon both their magnitude and their location within the molecules. If the dipole moment is not greater than about 1 Debye unit and is somewhat buried within the molecule, as it is in the cases of chloroform and ether, its effect may be mainly to alter the apparent solubility parameter, but larger or more exposed dipoles, such as hydroxyl, obviously introduce highly specific interactions which render equation 1 more or less inapplicable.

(*3*) Repulsive forces are simplified by assuming them to be of a uniform, radial, high-inverse-power type.

(*4*) The constants of attraction between unlike molecules are assumed to be the geometric mean between those for the like molecules (Berthelot relation). This is obviously violated whenever there are such interactions as hydrogen bonds, or electron donor–acceptor chemical reactions. These may be so small as to have eluded recognition by ordinary criteria and yet be sufficient to have considerable effects upon solubility relations. A striking illustration has recently been discovered by Benesi and Hildebrand (6) in the increasing solvation of iodine in benzene, toluene, xylene, and mesitylene, in that order, an apparently acid–base type of reaction that could scarcely have been predicted.

(*5*) The thermal agitation is assumed to be sufficient to overcome completely the segregating effect of unequal attractive forces. While this cannot be strictly true, there is a good deal of evidence that appreciable disturbance occurs only in the immediate neighborhood of a critical mixing point (4, 8, 11).

(*6*) Equation 1 does not take cognizance of changes of volume on mixing, although the formation of regular solutions is usually accompanied by expansion and increased entropy. The treatment can be refined by measuring and allowing for this expansion, but I propose in this critique deliberately to neglect this correction on the ground that our purpose is to predict the behavior of solutions from the properties of their pure components alone; if we must measure other properties of solutions in order to calculate solubility, we might as well measure the solubility itself. We can at least be reasonably sure of the sign and rough magnitude of the errors entailed by neglecting these volume changes. That we shall ever be able to calculate them for the wide variety of molecular types seems quite unlikely.

(*7*) The numerical values of the solubility parameters, δ, are often desired for temperatures far from those at which the heat of vaporization has been accurately determined, and the difference between the heat capacities of liquid and vapor, necessary for an extrapolation, are known only for comparatively few substances.

(*8*) In order to deal with the solubility of a solid, we refer its activity to the substance in the supercooled liquid form, to calculate which accurately we need not only the heat of fusion at the melting point but also the heat capacities of the solid and liquid forms. Even where these have both been measured, which is often not the case, the heat capacity of the liquid has to be extrapolated below the melting point.

With all the above sources of uncertainty, one might expect that equation 1, or, even where $v_1 \gg v_2$, equation 3, would yield only extremely rough figures for solubility; hence those who have become sufficiently conversant with both theory and data to judge have doubtless come, like myself, to feel both surprised and gratified to find how well it holds. This is well illustrated by the solubilities of iodine (table 2) in liquids of widely different solvent power, selected from a more extensive study by Benesi and Hildebrand (1). It will be noted that, although the solubility of iodine ranges from 0.0185 to 7.82 mole per cent, and its activity coefficient, a_2/x_2, in the saturated solutions ranges from 3.30 to 1400, these solubilities substituted in equation 1 yield δ_2 values for iodine agreeing within 3 per cent. Moreover, the dipoles present in chloroform, bromoform, and 1,2-dibromoethane and the large molal volume of normal perfluoroheptane appear to cause no serious disturbance. Numerous other cases of similarly good agreement have been presented.

We see, also, that the average of the δ-values for iodine calculated from its solubilities is a little larger than the value we obtain from its energy of vaporization per cubic centimeter. It is evident, therefore, that 14.1 would be a better figure than 13.6 to use for calculating its solubility in a new solvent, and this is not an isolated case. It seems worth while, therefore, to study the interrelations of a group of substances to see whether it is possible to derive a consistent set of δ-values from solubility data. It is, of course, desirable for this purpose to select solutions with large values of $\delta_2 - \delta_1$.

TABLE 2

Iodine solutions at 25°C.

SOLVENT	MOLECULAR VOLUME	MOLE PER CENT I_2	a_2^s/x_2	δ_1	δ_2
	cc.				
n-C_7F_{16}	227	0.0185	1400	5.7	14.2
$SiCl_4$	115.3	0.499	51.8	7.6	13.9
CCl_4	97.1	1.147	22.5	8.6	14.2
$TiCl_4$	110.5	2.15	12.0	9.0	14.1
$CHCl_3$	80.7	2.28	11.3	9.3	14.3
CS_2	60.6	5.46	4.73	9.9	14.1
$CHBr_3$	87.8	6.16	4.19	10.5	14.1
1,2-$C_2H_4Br_2$	86.6	7.82	3.30	10.4	14.1
Average					14.1
I_2	59.0	25.8	1.00	$(\Delta E^V/V)^{1/2} = 13.6$	

We have data for the solubilities at 25°C. in a more or less common group of solvents of the following solutes of comparatively high internal pressures, or δ-values: phosphorus (P_4), iodine (I_2), sulfur (S_8), stannic iodide, p-dibromobenzene, naphthalene, anthracene, and phenanthrene. The data are nearly all given, with references, in the third edition of *Solubility of Non-Electrolytes* (9) and need not be repeated here. The values of $\delta_2 - \delta_1$ in equation 1 calculated from these data are represented in figure 2. The vertical scales for the solutes have been displaced so as to make the line connecting the points for carbon

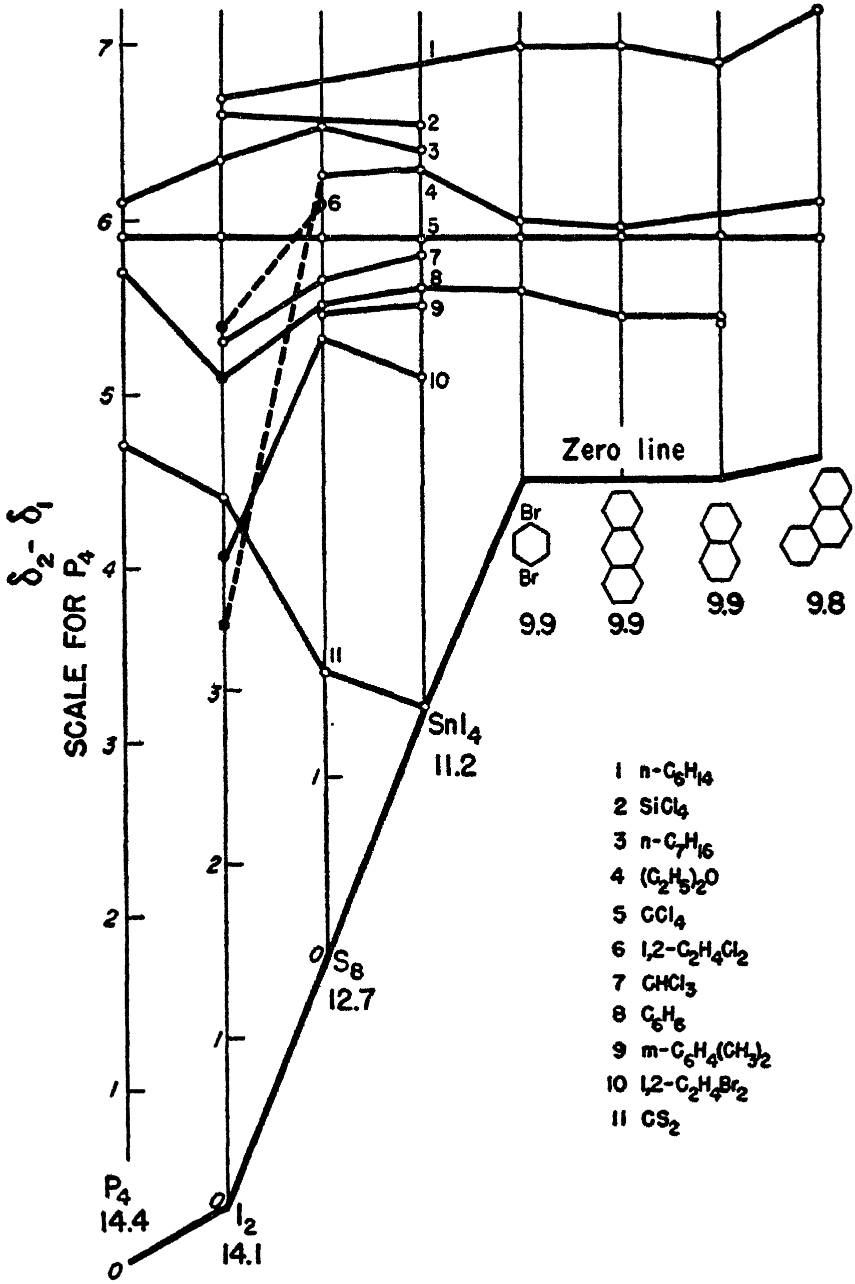


FIG. 2. A test of the internal consistence of solubility parameters

tetrachloride horizontal. If the equation applied rigidly to all the solutions represented, the other lines would also all be horizontal. The departures represent the various degrees of inadequacy of equation 1. The discrepancies are seen to be greatest for (*a*) iodine in ether, (*b*) 1,2-dichloroethane and 1,2-

dibromoethane, and (*c*) carbon disulfide. The first of these is explained by the evidence of strong solvation furnished by the brown color of the solution. The second pair of solvents have two strong dipoles each; hence some irregularity is to be expected. Carbon disulfide has the smallest molal volume of all the solvents represented, and some, at least, of the drop in its line in going from P_4 and I_2 to S_8 and SnI_4 may be explained by the increased difference in molal volumes. The line would be more nearly horizontal if the partial molal entropy of the solute were calculated by the aid of equation 2, instead of $-R \ln x_2$. The neglect of this factor is intentional.

The fluctuations of the lines for the various solvents from the horizontal seldom exceed 0.3 unit, if we exclude the cases of strong dipoles and solvation mentioned above, and we may regard this as determining the uncertainty in calculated values of solubility. In the case of iodine in carbon tetrachloride, for example, the measured values are: $v_1 = 97.1$ cc., $v_2 = 59.0$ cc., $a_2^s = 0.258$ (the activity of solid iodine referred to supercooled liquid), and $x_2 = 0.01147$. These give $\delta_2 - \delta_1 = 5.6$. Reversing the calculation and taking $\delta_2 - \delta_1 = 5.9$, one obtains $x_2 = 0.0094$, an error of 18 per cent. The logarithmic relation between x_2 and $(\delta_2 - \delta_1)^2$ naturally enhances the percentage error in calculating x_2.

It is noteworthy, next, that all the solubility data agree in assigning to chloroform a δ-value of about 9.0, less than that of benzene, 9.16, whereas its $(\Delta E^V/V)^{1/2} = 9.3$. It might be expected that its dipole moment would have some such effect, and it is our good fortune that it is so nearly the same with different solutes.

This study appears to me to justify the following general conclusions. Equation 1, or equation 2 if a volume ratio far from unity appears to make it worth the extra computation, applied to solutions in which specific interactions are absent, can yield figures for solubility of moderate reliability with uncertainties corresponding to 0.2–0.3 unit in δ-values. It can predict relative solubilities in different solvents with much smaller uncertainties. The direction and rough magnitude of departures from values which may be caused by association and solvation can often be fairly well estimated.

The quantitative limitations set forth in this paper are not so serious as to prevent the theory from being qualitatively very serviceable. We seek qualitative and relative solubility data far more often than exact figures. We seek the best or sometimes the poorest solvent for a certain solute. We seldom want to know a solubility to, say, 1 per cent and, indeed, we seldom control temperature or purity to a corresponding degree. If we do need a solubility to that accuracy we must rely upon measurement, better measurement, indeed, than many in the literature. All theory can do for us in that case is to select, out of the scores of solvents upon our shelves, the few likely to serve our purpose.

REFERENCES

(1) Benesi, H. A., and Hildebrand, J. H.: J. Am. Chem. Soc., in press.

(2) Hildebrand, J. H.: Proc. Natl. Acad. Sci. (U. S.) **13**, 267 (1927); J. Am. Chem. Soc. **51**, 66 (1929).

(3) Hildebrand, J. H.: J. Am. Chem. Soc. **59**, 794 (1937).
(4) Hildebrand, J. H.: Science **90**, 1 (1939).
(5) Hildebrand, J. H.: J. Chem. Phys. **15**, 225 (1947).
(6) Hildebrand, J. H., and Benesi, H. A.: J. Am. Chem. Soc. **70**, 2832 (1948).
(7) Hildebrand, J. H., and Gilman, T. S.: J. Chem. Phys. **15**, 229 (1947).
(8) Hildebrand, J. H., and Negishi, G. R.: J. Am. Chem. Soc. **59**, 339 (1937).
(9) Hildebrand, J. H., and Scott, Robert L.: *Solubility of Non-Electrolytes*, 3rd edition. Reinhold Publishing Corporation, New York (1949).
(10) Hildebrand, J. H., and Sweny, J. W.: J. Phys. Chem. **43**, 297 (1939).
(11) Kirkwood, J. G.: J. Phys. Chem. **43**, 97 (1939).
(12) London, F.: J. Chem. Phys. **46**, 305 (1942).

BINARY SOLUTIONS OF IMPERFECT LIQUIDS[1]

R. B. PARLIN AND H. EYRING

Department of Chemistry, University of Utah, Salt Lake City, Utah

Received October 25, 1948

A liquid mixture of two kinds of molecules is treated as composed of quasi-microcrystalline regions of twelve or less molecules. Interactions within these regions are explicitly represented by a Boltzmann factor with specific energies for each kind of molecular pair. Interactions between molecules in different crystalline regions are represented by a Boltzmann factor involving an averaged energy term. The partition function of the mixture is obtained by summing over all configurations, weighting each term by the appropriate Boltzmann factors. In the case of approximately equal-sized molecules, explicit expressions are obtained for the properties of the solution, notably the consolute temperature, in terms of the coördination number and the interaction energies. For mixtures of molecules of widely different sizes, the corresponding equations are derived but explicit results have not yet been obtained.

INTRODUCTION

A number of treatments have been proposed for the problem of the interactions between molecules in the liquid state (4, 5), most of which can be shown to give the essential properties of such systems in a more or less quantitative manner. Generally, the most profitable approach has been that of considering the liquid to be a distorted crystal, in which the long-range order has been replaced by smaller regions within which a lattice array of molecules is at least momentarily maintained, but between which no order is observed.

To treat this ensemble of cells, it is necessary to approximate the configurational potential energy integral in the partition function (3) in some appropriate manner. In this paper, an adaptation of the order-disorder theory originally due to Bethe (1) and later applied by Peierls (7) to the adsorption of gases on surfaces and by Cernuschi and Eyring (2; see also 6) to the theory of condensation will be used. The theory is applied directly to a simple model of approximately equal-sized molecules, and the direction of extension to more realistic systems is indicated.

THEORY

It will be assumed (*1*) that the molecules of the liquid occupy positions of a lattice (although the lattice will not be constrained to maintain any but topological identity) and each molecule will be supposed to have a number Z of nearest neighbors; (*2*) that when to a molecule of kind i a nearest neighbor of kind j is added to the lattice there is a resultant enhancement of the probability of i being there owing to the attraction between the molecules; (*3*) that all interactions of a molecule in the first neighboring shell with any but the central molecule will be represented by an averaged energy.

---

[1] Presented at the Symposium on Thermodynamics and Molecular Structure of Solutions, which was held under the auspices of the Division of Physical and Inorganic Chemistry at the 114th Meeting of the American Chemical Society, Portland, Oregon, September 13 and 14, 1948.

Consider then such a mixture of molecules (designated as kind 1 and kind 2), and let the ratio of the partition function for a molecule in the liquid to that in the gas be represented by f_1 and f_2. In view of the greater deviation from ideality for the liquid, the gas will for the time being be considered a perfect gas.

Suppose a given lattice site is occupied by a molecule of kind 1; several possible configurations may then arise (figure 1).

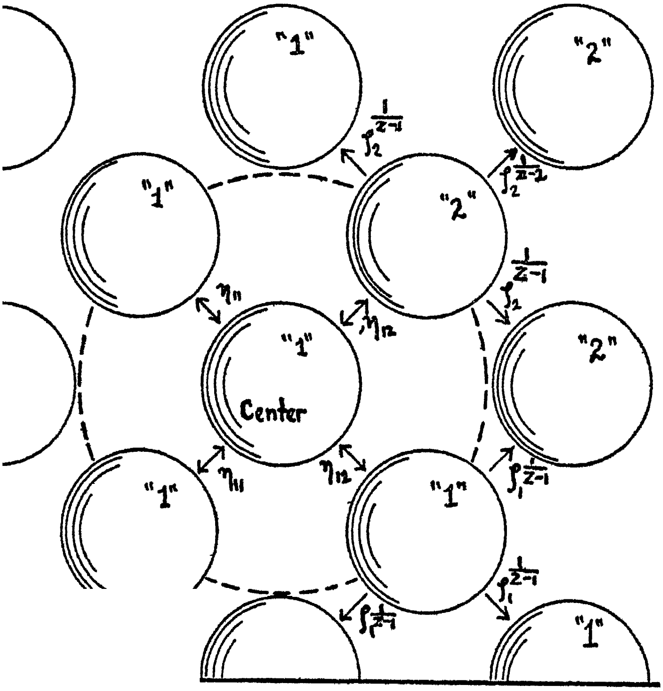


FIG. 1. Representation of a mixture of molecules of kinds 1 and 2 as a plane lattice

(*1*) If all the positions in the first shell are occupied by similar molecules, the partition function for this configuration will be given by the expression

$$f_1 \cdot f_1\zeta_1\eta_{11} \cdot f_1\zeta_1\eta_{11} \cdot f_1\zeta_1\eta_{11} \ldots . = f_1 \cdot (f_1\zeta_1\eta_{11})^Z \qquad (1)$$

where the factor ζ_1 corrects for all interactions of a shell molecule except the interaction with the central molecule, as assumed previously, and η_{11} provides for this omitted interaction.

(*2*) If all of the positions in the first shell except one are of kind 1 and the Z^{th} is of kind 2, the partition function will be

$$Z \cdot f_1 \cdot f_1\zeta_1\eta_{11} \cdot f_1\zeta_1\eta_{11} \ldots . \cdot f_2\zeta_2\eta_{12} = Z \cdot f_1 \cdot (f_1\zeta_1\eta_{11})^{Z-1} \cdot (f_2\zeta_2\eta_{12}) \qquad (2)$$

where the factor Z acknowledges that such a configuration may occur in Z ways. Continuing in this wise, it may readily be verified that the complete expression is the sum of the terms in the binomial expansion thus:

$$f_1[f_1\zeta_1\eta_{11} + f_2\zeta_2\eta_{12}]^Z \tag{3}$$

The partition function, one sees, is the sum of the weighted terms for each configuration. Clearly the analogous result for the case in which the central site is occupied by a molecule of kind 2 will be

$$f_2[f_1\zeta_1\eta_{12} + f_2\zeta_2\eta_{22}]^Z \tag{4}$$

The ratio of the two expressions (3 and 4) will be equal to the ratio of the probabilities of the central point being occupied by molecule 1 to that of its being occupied by molecule 2, that is, to the mole ratio $\theta/1-\theta$:

$$\frac{\theta}{1-\theta} = \frac{f_1}{f_2} \cdot \frac{[f_1\zeta_1\eta_{11} + f_2\zeta_2\eta_{12}]^Z}{[f_1\zeta_1\eta_{12} + f_2\zeta_2\eta_{22}]^Z} \tag{5}$$

On the other hand, if one counts the configurations in this manner, the mole ratio $\theta/1-\theta$ must also be given by the expression

$$\frac{\phi}{1-\theta} = \frac{f_1\zeta_1}{f_2\zeta_2} \cdot \frac{[f_1\zeta_1\eta_{11} + f_2\zeta_2\eta_{12}]}{[f_1\zeta_1\eta_{12} + f_2\zeta_2\eta_{22}]} \tag{6}$$

where attention is now focussed upon a molecule in the shell of neighbors. Here there are two configurations if the particular neighboring molecule is of kind 1, since the central molecule may be of either kind. In one case the partition function will be $f_1\zeta_1 \cdot f_1\zeta_1\eta_{11}$, while if the central molecule is of kind 2 one will have $f_1\zeta_1 \cdot f_2\zeta_2\eta_{12}$. The partition function for the case of the type 1 shell molecule will then be the sum of these expressions. A similar approach gives the result for the case of a molecule of kind 2 in a shell position, and the ratio of these two sums is then the mole ratio as indicated in equation 6. Since the distinction between a "central" position and a "neighboring" position is a purely formal one, the two partition function ratios clearly must be identical.

It is convenient, for purposes of computation, to make the substitutions

$$\begin{aligned} f &= f_1/f_2 & \eta &= \eta_{12}/\eta_{22} \\ x &= f_1\zeta_1/f_2\zeta_2 & \mu &= \eta_{11}/\eta_{22} \end{aligned} \tag{7}$$

when equations 5 and 6 become, upon rearrangement

$$f = x\left(\frac{1+\eta x}{\eta + \mu x}\right)^{z-1} \tag{8}$$

$$\frac{\theta}{1-\theta} = x\left(\frac{\eta + \mu x}{1+\eta x}\right) \tag{9}$$

One may identify the function f with the ratio $p_1/p_1^0/p_2/p_1^0$, where p_1 is the partial pressure of component 1 and p_1^0 is the vapor pressure of the pure liquid 1. Since only for a perfect liquid mixture is the total pressure a linear function of the composition, it is necessary to obtain a third relation which will permit evaluation of the individual partial pressures, rather than their ratio alone. This desired relation is obtained from the Gibbs equation

$$\theta\left(\frac{\partial\mu_1}{\partial\theta}\right)_{T,P} = (1-\theta)\left(\frac{\partial\mu_2}{\partial(1-\theta)}\right)_{T,P} \tag{10}$$

(where θ represents the mole fraction of component 1, and μ_1 the partial molar free energy of the same substance in the solution). Making use of the definition of f and remembering that the gases are considered ideal, whence

$$\mu_1 = RT\ln p_1/p_1^0 \qquad \mu_2 = RT\ln p_2/p_2^0 \tag{11}$$

one obtains the result that

$$\mathrm{d}\ln\frac{p_2}{p_2^0} + \frac{\theta}{f}\,\mathrm{d}f = 0 \tag{12}$$

Equation 12 may be integrated in the form

$$\frac{p_2}{p_2^0} = \exp\left(-\int_0^f \frac{\theta}{f}\,\mathrm{d}f\right), \qquad \frac{p_1}{p_1^0} = f\cdot\frac{p_2}{p_1^0} \tag{13}$$

or, for computational purposes, expressed in terms of the variable x as

$$\ln\frac{p_2}{p_2^0} = -\eta\int_0^x\left[1 - \frac{(Z-2)\left(\frac{\mu}{\eta}-\eta\right)x}{\mu x^2 + 2\eta x + 1}\right]\mathrm{d}x \tag{14}$$

Integrating equation 14 one finds two solutions, depending on the relative values of the constants. For $\eta^2 > \mu$

$$\ln\frac{p_2}{p_2^0} = \eta x + \frac{(Z-2)(\eta^2-\mu)}{2\mu}\ln\left[\mu x^2 + 2\eta x + 1\right] - \frac{(Z-2)\eta(\eta^2-\mu)^{1/2}}{2\mu}\ln\left\{\frac{\eta+(\eta^2-\mu)^{1/2}}{\eta-(\eta^2-\mu)^{1/2}}\cdot\frac{\mu x+\eta-(\eta^2-\mu)^{1/2}}{\mu x+\eta+(\eta^2-\mu)^{1/2}}\right\} \tag{15}$$

while in the case $\eta^2 < \mu$,

$$\ln\frac{p_2}{p_2^0} = \eta x - \frac{(Z-2)(\mu-\eta^2)}{2\mu}\ln\left[\mu x^2 + 2\eta x + 1\right) + \frac{(Z-2)\eta(\mu-\eta^2)^{1/2}}{\mu}\left\{\tan^{-1}\frac{\mu x+\eta}{(\mu-\eta^2)^{1/2}} - \tan^{-1}\frac{\eta}{(\mu-\eta^2)^{1/2}}\right\} \tag{16}$$

The two forms (equations 15 and 16) of the integral express the fact that two possibilities may arise when two liquids are mixed at a given temperature. Equation 15 gives the result for $\ln p_2/p_2^0$ when there is no maximum in the p_2–θ relation. Equation 16 permits of such a maximum, and hence expresses the result for a two-phase region (see figure 2).

To examine this point more closely, one may observe that the function θ, the mole fraction of component 1, has the derivative

$$\frac{\mathrm{d}\theta}{\mathrm{d}x} = \frac{\eta\mu x^2 + 2\mu x + 1}{[\eta\mu x^2 + 2\eta\mu x + \mu]^2} \tag{17}$$

so that θ is a monotonic function of the variable x (μ and η being intrinsically positive) while the derivative of f,

$$\frac{\mathrm{d}f}{\mathrm{d}x} = \left(\frac{1 + \eta x}{\eta + \mu x}\right)^{Z-1} \left[1 - \frac{(Z-1)(\mu - \eta^2)x}{(1 + \eta x)(\eta + \mu x)}\right] \tag{18}$$

shows that f has a maximum *and* a minimum if

$$\eta^2 < \left(\frac{Z - 2}{Z}\right)\mu \tag{19}$$

the corresponding values of f being determined from equation 8 and

$$x^* = \frac{(Z-2)\mu - Z\eta^2}{2\eta\mu}\left[1 \pm \sqrt{1 - \frac{4\eta}{[(Z-2)\mu - Z\eta^2]^2}}\right] \tag{20}$$

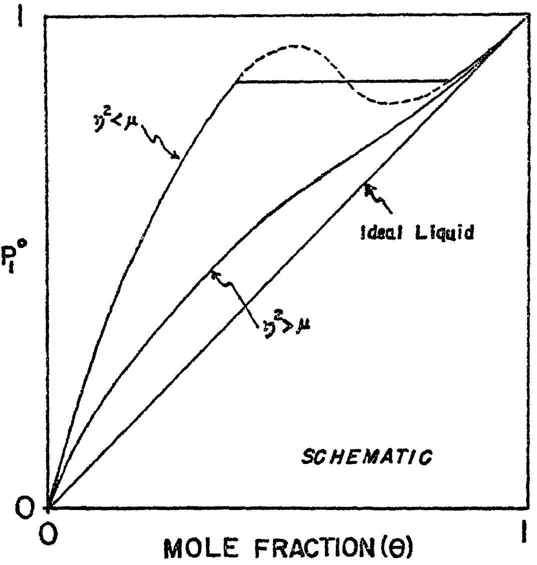


FIG. 2. Vapor pressure–composition curve for a binary solution

Now one may write, for $x = x^*$, the expression

$$\frac{\mu}{\eta^2} = \frac{\left(\theta + \frac{1}{Z-2}\right)\left(1 - \theta + \frac{1}{Z-2}\right)}{\theta(1-\theta)} \tag{21}$$

Equation 21 gives the relation between the composition of the pairs of equilibrium solutions and the temperature, since the η, μ may be expressed as Boltzmann factors of the form

$$\eta = e^{(\epsilon_{12} - \epsilon_{22})/kT} \qquad \mu = e^{(\epsilon_{11} - \epsilon_{22})/kT} \tag{22}$$

Using equations 22 one obtains, setting

$$\epsilon = \frac{1}{2}(\epsilon_{11} + \epsilon_{22}) - \epsilon_{12} \tag{23}$$

the result

$$T = 2\epsilon \quad k \ln \frac{\left(\theta + \frac{1}{Z - 2}\right)\left(1 - \theta + \frac{1}{Z - 2}\right)}{\theta(1 - \theta)} \tag{24}$$

The temperature is maximum at $\theta = 1/2$, and its maximum value is

$$T_{\text{consolute}} = \frac{\epsilon}{k \ln \frac{Z}{Z - 2}} \tag{25}$$

It is interesting to note that if the coördination number Z were to change suddenly with temperature, i.e., to drop from 12 to 4 with a few degrees increase in temperature, it would be possible to be above the consolute temperature at low temperature (and low coördination) and to be below the mixing temperature at high temperatures (and high coördination). Such a situation would give rise, albeit crudely, to the phenomenon of a lower consolute point. One may also readily see that under certain circumstances, the upper consolute temperature may be impossible of attainment, at least under conditions of normal pressures. In fact, unless

$$T_{\text{melting}} < T_{\text{consolute}} < T_{\text{boiling}} \tag{26}$$

the liquid pairs will either be completely miscible or will be partially soluble over the whole range of temperatures of the liquid state. Now approximately, one may write

$$T_{\text{boiling}} \approx \frac{\Delta H_v}{21} \tag{27}$$

by Trouton's rule, and

$$\Delta H_v \approx 6(\epsilon_{11}\theta + (1 - \theta)\epsilon_{22}) \tag{28}$$

since the energy of about six bonds in the liquid state must be broken to vaporize a molecule (assuming twelve neighbors). Then

$$T_{\text{boiling}} \approx 0.29(\epsilon_{11}\theta + (1 - \theta)\epsilon_{22}) \tag{29}$$

Since $\theta = 1/2$ for the consolute composition we have for this composition:

$$T_{\text{boiling}} \approx 0.15(\epsilon_{11} + \epsilon_{22}) \tag{30}$$

Dividing through by equation 25 gives:

$$\frac{T_{\text{boiling}}}{T_{\text{consolute}}} \approx 0.15 \frac{\epsilon_{11} + \epsilon_{22}}{\epsilon} k \ln \frac{Z}{Z - 2} \tag{31}$$

and $0.15 \frac{\epsilon_{11} + \epsilon_{22}}{\epsilon} k \ln \frac{Z}{Z - 2}$ must be greater than unity for the consolute temperature to appear in the liquid region. If we write:

$$T_{\text{melting}} = \alpha T_{\text{boiling}} \tag{32}$$

whence

$$\frac{T_{\text{melting}}}{T_{\text{consolute}}} \approx 0.15k\alpha\left(\frac{\epsilon_{11}+\epsilon_{22}}{\epsilon}\right)\ln\frac{Z}{Z-2} \tag{33}$$

If we make the estimation $\alpha = 0.73$ as in the case of water we obtain:

$$\frac{T_{\text{melting}}}{T_{\text{consolute}}} = 0.11k\left(\frac{\epsilon_{11}+\epsilon_{22}}{\epsilon}\right)\ln\frac{Z}{Z-} \tag{34}$$

Referring again to equation 25, it is apparent that since ϵ (as defined by equation 23) may be negative, there should exist systems for which no consolute temperature may be found, even in principle, above the absolute zero. That is, if the mean of the interactions between like pairs of molecules should fall below the interaction between unlike pairs, the two liquids should be miscible in all proportions at any temperature whatever, provided of course that such temperatures may be obtained without fusion or vaporization. In other cases, with sufficiently positive ϵ, a consolute or critical mixing temperature should be realizable, perhaps under high pressures.

To examine the applicability of equations 3, 9, and 14 to experimental data, it is convenient to plot the ratio of partial pressures against the ratio of mole fractions (or the mole ratio). From equations 8 and 9 one may note that

$$f = \left(\frac{1+\eta x}{\eta+\mu x}\right)^{Z}\frac{\theta}{1-\theta} \tag{35}$$

and the relation is linear for f plotted against $\theta/1-\theta$ in the region of small x (hence small θ). The slope of this limiting straight line will then be η^{-Z}. In addition, when $f = \frac{\theta}{1-\theta}$, $\frac{1+\eta x}{\eta+\mu x} = 1$, and

$$\mu - \eta$$

hence

$$f = \frac{\theta}{1-\theta} = \frac{1-\eta}{\mu-\eta} \tag{36}$$

at the point at which the curve crosses the ideal solution line, $f = \theta/1-\theta$ (figure 3).

The form of such experimental curves may be seen in the model curves of figure 4, drawn for the set of constants,

$$\begin{aligned} Z &= 12 \\ \epsilon_{12}-\epsilon_{22} &= -\ 84.0 \text{ cal./mole} \\ \epsilon_{11}-\epsilon_{22} &= 209.6 \text{ cal./mole} \end{aligned} \tag{37}$$

With this choice of constants, the consolute temperature, or critical solution temperature, will be about 400°K.

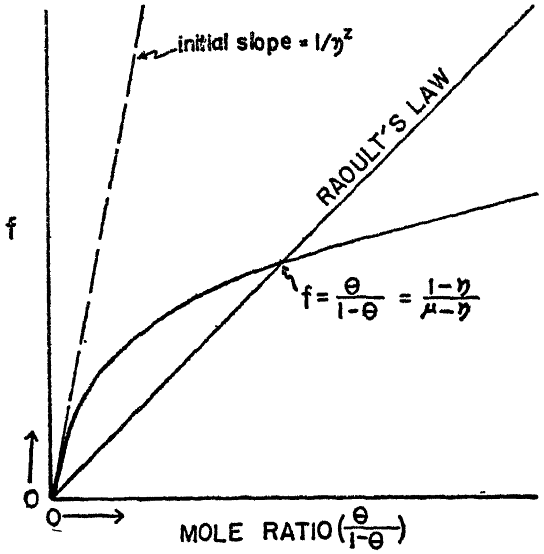


FIG. 3. Plot of partial pressure ratio *vs.* mole ratio for determining constants of equation

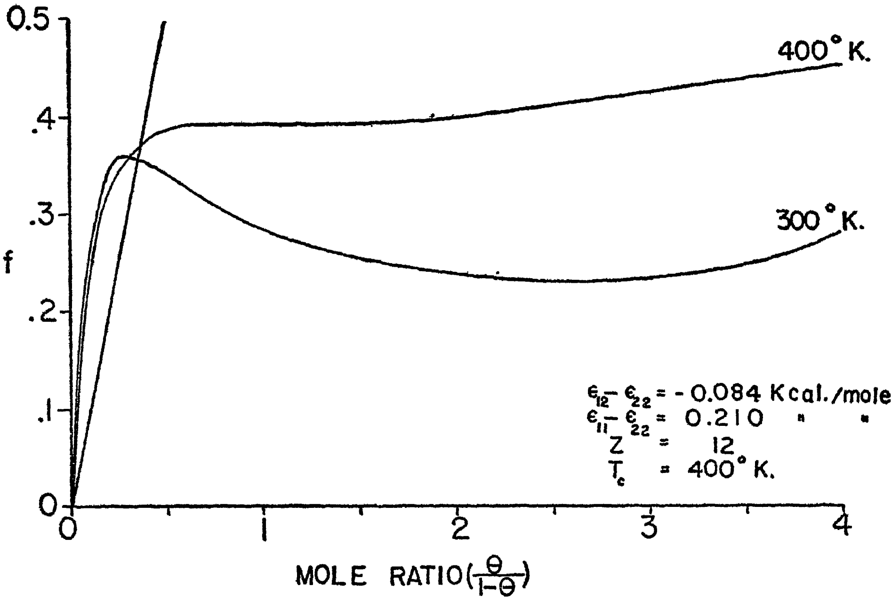


FIG. 4. Model curve

Finally, one may readily test equations 24 and 25. Equation 24 may be arranged in the form,

$$\frac{T}{T_{\text{consolute}}} = 2\,\frac{\ln Z/Z - 2}{\ln \dfrac{\left(\theta + \dfrac{1}{Z-2}\right)\left(1 - \theta + \dfrac{1}{Z-2}\right)}{\theta(1-\theta)}} \tag{38}$$

which is a one-parameter family of curves for $T/T_{\text{consolute}}$ in terms of θ. Figure 5 gives this relation in graphical form for selected values of the coördination number Z. From equation 29 it should be possible to determine Z, and from the relation

$$\epsilon = kT_{\text{consolute}} \ln \frac{Z}{Z - 2} \tag{39}$$

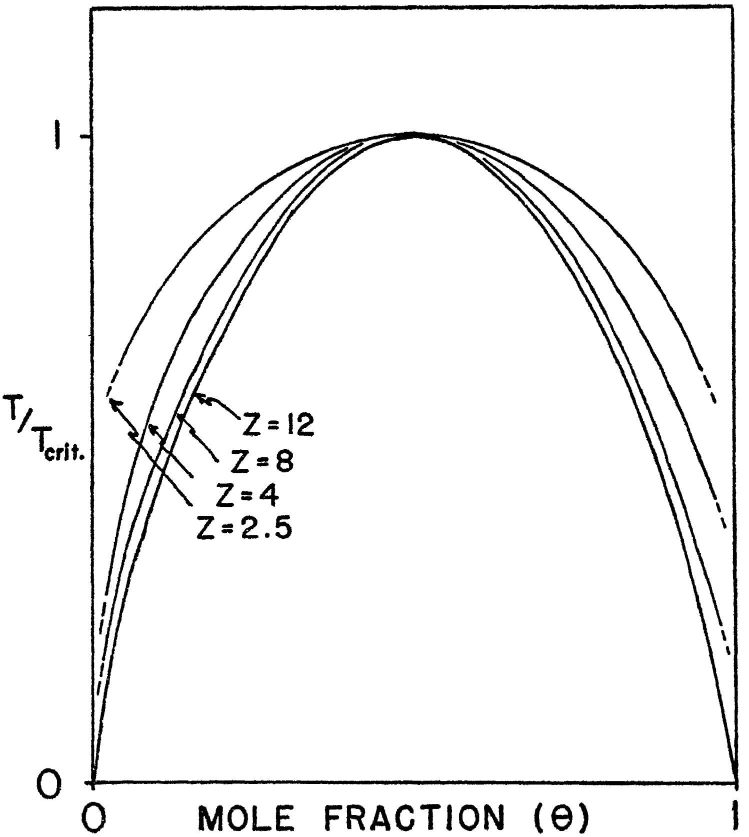


Fig. 5. Temperature (in terms of consolute temperature) as a function of solubility for various values of the coördination number Z.

one may then obtain the value of ϵ, the difference between the mean of the interaction energies between like molecules and the interaction energy for unlike molecules. It should be noted, of course, that expression 25 relates the temperature not strictly to the solubility, but to the value of θ corresponding to the maximum of the vapor pressure–composition curve. Clearly, the actual composition at the saturation point will be somewhat less than the θ of equation 29 on one branch of the curve and somewhat greater on the other. In general, however, there will be some curve of this description, and the principles involved will be the same.

Consider now the case of a mixture of two kinds of molecules, of which one sort occupies two lattice positions while the other occupies but one (figure 6). The double molecules will be designated by the subscript 2, the single particles by 1. To obtain the partition function for a molecule in the field of its nearest neighbors one writes the sum of all possible combinations; if the central molecule is of kind 2, there will be a contribution f_2 to the partition function in the absence of neighbors. If all of the $Z - 1$ nearest lattice points contain type 1 particles, the contribution is

$$(f_1\zeta_1\eta_{12})^{Z-1}\cdot f_2 \tag{40}$$

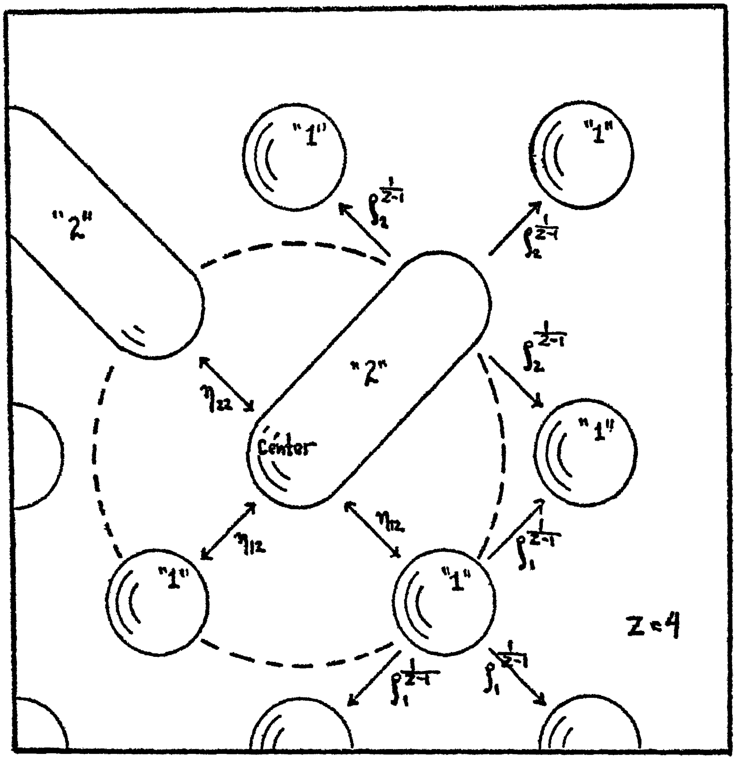


FIG. 6. Representation of a mixture of single (1) and double (2) molecules as a plane lattice

Continuing in this fashion, the total sum will be found to be

$$f_2[f_2\zeta_2\eta_{22} + f_1\zeta_1\eta_{12}]^{Z-1} \tag{41}$$

Similarly, if the central point is occupied by a molecule of kind 1, the corresponding expression will be

$$f_1[f_2\zeta_2\eta_{12} + f_1\zeta_1\eta_{11}]^{Z} \tag{42}$$

In the same way, if one considers a molecule of kind 2 in the first shell, its partition function will be

$$f_2\zeta_2\,[f_2\zeta_2\eta_{22} + f_1\zeta_1\eta_{12}]^{(Z-1)/Z} \tag{43}$$

while that for a molecule of kind 1 will be

$$f_1\zeta_1[f_2\zeta_2\eta_{12} + f_1\zeta_1\eta_{11}] \tag{44}$$

Now the ratio of expression 27 to 28, as of expression 29 to 30, must be the ratio of the probabilities of finding an arbitrary molecule to be of kind 1 or kind 2, hence the mole ratio:

$$\frac{\theta}{1-\theta} = \frac{f_2}{f_1}\,\frac{[f_2\zeta_2\eta_{22} + f_1\zeta_1\eta_{12}]^{z-1}}{[f_2\zeta_2\eta_{12} + f_1\zeta_1\eta_{11}]^{z}} = \frac{f_2\zeta_2}{f_1\zeta_1}\,\frac{[f_2\zeta_2\eta_{22} + f_1\zeta_1\eta_{12}]^{(z-1)/z}}{[f_2\zeta_2\eta_{12} + f_1\zeta_1\eta_{11}]} \tag{45}$$

Equations 31 may be written in the form

$$\frac{\theta}{1-\theta} = \frac{f_2}{f_1}\cdot\frac{(f_2\zeta_2)^{z-1}}{(f_1\zeta_1)^{z}}\cdot\frac{\left(1 + \eta\dfrac{f_1\zeta_1}{f_2\zeta_2}\right)^{z-1}}{\left(\eta + \mu\dfrac{f_1\zeta_1}{f_2\zeta_2}\right)^{z}} \tag{46}$$

$$= \left[\frac{(f_2\zeta_2)^{z-1}}{(f_1\zeta_1)^{z}}\right]^{1/z}\cdot\frac{\left(1 + \eta\dfrac{f_1\zeta_1}{f_2\zeta_2}\right)^{(z-1)/z}}{\left(\eta + \mu\dfrac{f_1\zeta_1}{f_2\zeta_2}\right)} \tag{47}$$

Since in this case the expressions are not homogeneous in $f_1\zeta_1$ and $f_2\zeta_2$, the simple expressions analogous to expressions 6 and 7 are not possible, and a second implicit variable $(f_2\zeta_2)^{z-1}/(f_1\zeta_1)^z$ appears. However there is still a third equation, the Gibbs equation, which has again the form

$$\frac{p_2}{p_1^0} = \exp\left\{-\int_{\infty}^{f_2/f_1}\theta\,\mathrm{d}\ln\frac{f_2}{f_1}\right\} = \frac{p_1/p_1^0}{f} \tag{48}$$

and in principle it should be possible to calculate the equilibrium curves of partial pressure *vs.* composition. In practice, the expressions have been found too unwieldy for numerical computation as yet. One may note that for large values of coördination number, Z,

$$\frac{(f_2\zeta_2)^{z-1}}{(f_1\zeta_1)^{z}} \text{ reduces approximately to } \frac{f_2\zeta_2}{f_1\zeta_1}$$

or in terms of the previous notation to x^z.

A number of difficulties arise in the manipulation of the equations for any case except the simple one of molecules of equal size. In principle, however, it should be possible to treat almost any kind of mixture of arbitrary molecules to the point of setting up the equations. In particular, it is noteworthy that special kinds of interactions such as forces between dipoles may readily be included, as it is only necessary to identify such special forces with one of the η_{ij}, the energy for which will be determined by the nature of the case.

Further, it is possible, and may in some cases be essential, to consider the various ϵ_{ij} not as energies (heats) but rather as (Helmholtz) free energies: thus if the ϵ_{ij} are considered to be linear functions of the temperature, one may write for ϵ,

$$\epsilon = \frac{1}{2}(\epsilon_{11} + \epsilon_{22}) - \epsilon_{12} = E - TS \tag{49}$$

where the terms E and TS represent the temperature-independent (energy) and temperature-dependent (entropy) parts, respectively. The counterpart to equation 25 will then be

$$T_{\text{consolute}} = \frac{E}{S + k \ln \frac{Z}{Z - 2}} \tag{50}$$

for the consolute temperature. It is apparent that if the term S in the denominator were to be larger than the coördination number term, the consolute temperature would be sensibly independent of the coördination.

Finally, one may remember the empirical fact that the use of volume fractions in place of mole fractions in many cases removes the anomalies found in some of the properties of liquids which arise from the unequal size of the molecules. It is tempting to suppose that the variation in the consequences of equations 46, 47, and 48 from those of equations 8, 9, and 13 which arise from just such a source may be reproduced by using the simpler theory with the mole fraction terms replaced by volume fractions.

REFERENCES

(1) BETHE, H.: Proc. Roy. Soc. (London) **A150,** 552 (1935).

(2) CERNUSCHI, F., AND EYRING, H.: J. Chem. Phys. **7,** 547 (1939).

(3) FOWLER, R. H., AND GUGGENHEIM, E. A.: *Statistical Thermodynamics*, p. 322. The Macmillan Company, New York (1939).

(4) Reference, 3, pp. 350–69.

(5) KIRKWOOD, J. G.: J. Phys. Chem. **43,** 97 (1939).

(6) LENNARD-JONES, J. E., AND DEVONSHIRE, A. F.: Proc. Roy. Soc. (London) **A169,** 317 (1939).

(7) PEIERLS, R.: Proc. Cambridge Phil. Soc. **32,** 477 (1936).

THE DIELECTRIC CONSTANTS OF HYDROGEN-BONDED SUBSTANCES[1]

WENDELL M. LATIMER

Department of Chemistry, University of California, Berkeley, California

Received October 25, 1948

Evidence is presented to show that the abnormally high values for the dielectric constant of many hydrogen-bonded substances are probably due to the "atomic polarization" of the proton in the bond. A mechanism is suggested for the orientation of the dipoles in ice which depends upon the jumping of protons over the potential barrier in the hydrogen bond and not upon the rotation of the water molecules. The calculated activation energy for the process is in agreement with the experimental values obtained by Wentsch.

My discussion will be confined largely to two subjects: (*1*) the general effect upon the dielectric constant of small "atomic polarizations" due to proton displacements and (*2*) a mechanism for the orientation of the dipoles in ice which does not require a rotation of the molecules.

As an introduction, a brief review will be given of the status of the general problem of the dielectric properties of liquids. The Clausius–Mosotti equation, relating the dielectric constant and the distortion polarization, and the extension by Debye (1) to include the effect of permanent dipole account most satisfactorily for the behavior of gases and dilute solutions of polar molecules in non-polar solvents:

$$\frac{\epsilon - 1}{\epsilon + 2} = \frac{4\pi N}{3}\left(\alpha + \frac{\mu}{3kT}\right) \tag{1}$$

where ϵ is the dielectric constant, α the distortion polarization, N the number of molecules per cubic centimeter, and μ the permanent moment. Its application to liquids fails for the same general reason that the perfect gas law cannot be used to calculate pressure–volume relations in liquids. The problem of relating the dielectric constant of a liquid to the properties of the individual molecules is obviously highly complicated and doubtless not subject to exact formulation, but substantial progress has been made toward an approximate solution.

Wyman (15) observed that for non-associated liquids the dielectric constant is linearly proportional to the polarization per cubic centimeter calculated from the moment of the gas molecule and gave the equation:

$$(\epsilon + 1) = AP \tag{2}$$

The data for ethyl formate and acetone in table 1 indicate the order of agreement which he found, with A equal to 8.5.

---

[1] Presented at the Symposium on Thermodynamics and Molecular Structure of Solutions, which was held under the auspices of the Division of Physical and Inorganic Chemistry at the 114th Meeting of the American Chemical Society, Portland, Oregon, September 13 and 14, 1948.

Equation 2 does not hold for hydrogen-bonded liquids, as is illustrated by the data for water and formic acid.

Onsager (11) recognized the importance of Wyman's contribution. He wrote: "Wyman's relation implies quite simply that the polarization of the dielectric (by orientation) involves practically no change of the internal energy .... Such a discovery belies the pessimistic expectations of the 'association theory' and indicates that our theoretical understanding of polar liquids is capable of great advancement."

Onsager treated the molecule as a cavity of dielectric constant equal to the square of the optical refractive index of the liquid surrounded by a continuum of liquid of uniform dielectric constant. He introduced the concept of a reaction field which acts upon the dipole as a result of the electric displacements induced by its own presence and obtained the equation:

$$\frac{(\epsilon - n^2)(2\epsilon + n^2)}{\epsilon(n^2 + 2)} = \frac{4\pi N\mu^2}{9kT} \tag{3}$$

TABLE 1

Comparison of Wyman's calculations with experiment

COMPOUND	P (FROM GAS)	ϵ_{exptl}	ϵ_{calcd}
Ethyl formate	1.21	8.6	9.4
Acetone	2.62	21.5	21.1
Water	4.24	78.5	35.0
Formic acid	1.08	62	8.2

For high values of ϵ this reduces approximately to the Wyman equation with

$$A = \frac{(n^2 + 2)^2}{2} \tag{4}$$

and the value of A equal to 8.5 corresponds to a reasonable average value of 1.46 for n. The Onsager equation also fails to account for the high dielectric constant of hydrogen-bonded liquids and would require that μ for the water molecule in the liquid be around 3×10^{-18}. Near the end of his paper Onsager makes the remark, "Incidentally our 'refractive index' should include the 'atomic polarization' due to elastic displacement of the atomic nuclei by the electric field." It is apparent from equations 2 and 4 that even small atomic polarizations will have a significant effect upon ϵ, since it enters the constant A. This point will be amplified later in more detail.

Kirkwood (7) has attacked the problem of the hydrogen-bonded liquids by assuming hindered rotation of the dipole due to short-range intermolecular forces. His treatment leads to the equation

$$\frac{(\epsilon - 1)(2\epsilon + 1)}{9\epsilon} \cdot V = \frac{4\pi N}{3}\left[\alpha + \frac{g\mu^2}{3kT}\right] \tag{5}$$

where V = molar volume, α = optical polarizability, μ = the dipole moment of the liquid, and g is given by the expression

$$g = [1 + Z\,(\cos\gamma)_{\text{av}}] \tag{6}$$

where Z is the average number of nearest neighbors and $(\cos\gamma)_{\text{av}}$ is the average value of the cosine of the angle between the given dipole and one of its neighbors.

The effectiveness of Kirkwood's treatment is indicated by the calculated and observed values of ϵ for water and the aliphatic alcohols summarized in tables 2 and 3. In making these calculations the dipole moment μ of equation 4 was calculated from the Onsager expression for the relationship between μ and μ_0 of the molecule as a gas:

TABLE 2

Dielectric constant of water

t	ϵ_{calcd}	ϵ_{obsd}
°C.		
0	84.2	88.0
25	78.2	78.5
62	72.5	66.1
83	67.5	59.9

TABLE 3

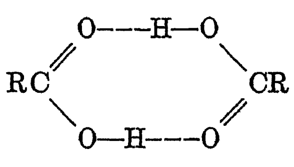
Dielectric constant of the aliphatic alcohols

ALCOHOL	ϵ_{calcd}	ϵ_{obsd}
Methyl alcohol	29.2	32.8
Ethyl alcohol	21.3	24.6
n-Propyl alcohol	17.3	19.5
n-Butyl alcohol	14.6	18.0
n-Amyl alcohol	12.6	15.8

$$\mu = \frac{2\epsilon + 1}{2\epsilon + n^2} \cdot \frac{n^2 + 2}{}\,\mu_0 \tag{7}$$

ATOMIC POLARIZATION IN HYDROGEN-BONDED LIQUIDS

There is considerable evidence for a very appreciable contribution to the dielectric constant of hydrogen-bonded liquids by the "atomic polarization" of the proton in the hydrogen bond. Hobbs, Gross, and coworkers (9, 12) have studied the apparent moments of the carboxylic acids in organic solvents. The acids were present as equilibrium mixtures of the monomer and the symmetrical dimer,



Evaluation of the parameters by successive approximations gave the equilibrium constant and the apparent moments of the monomer and dimer. From their data they concluded that the symmetrical dimer had an atomic polarization of about 20 cc.

Coop, Davidson, and Sutton (2) found a similar moment in the formic acid dimer in the vapor phase. They showed that the moment was independent of the temperature and therefore not due to an orientation type of polarization.

Recently Collie, Hasted, and Ritson (1) have reported an extensive study of the dielectric constant of light and heavy water as a function of frequency. They analyzed their data in terms of the equation,

$$\epsilon - \epsilon_0 = \frac{\epsilon_s - \epsilon_0}{1 + j\omega\tau} \tag{8}$$

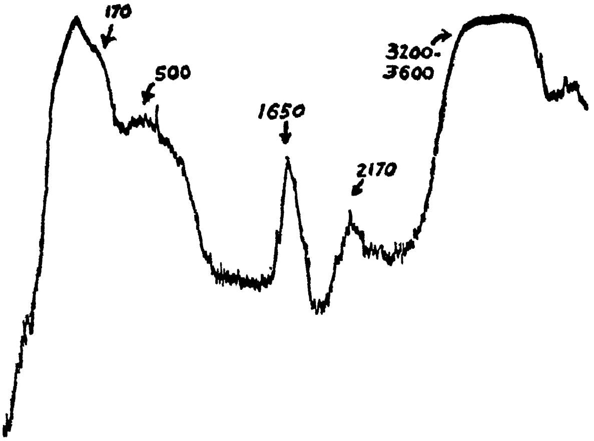


FIG. 1. Raman spectrum of water (from Hibben)

where τ is the relaxation time, ϵ the complex dielectric constant at angular frequency ω, ϵ_0 the electronic and atomic polarization, and ϵ_s the static dielectric constant. At constant temperatures equation 8 gives an accurate representation of their results with a constant value of τ. They find $\epsilon_0 = 5.5$ for both light and heavy water, a value which is considerably larger than the square of the index of refraction, and conclude that the atomic polarization in water is much higher than is generally assumed.

It seems likely that an atomic polarization of hydrogen-bonded substances is to be associated with a low-frequency vibration of the proton and it is therefore of interest to examine the infrared spectrum for such frequencies.

Hibben (5) has given the Raman spectrum of water reproduced in figure 1. The bands at 500, 1650, and 3200–3400 cm.$^{-1}$ are hydrogen vibrations, since they are shifted by approximately $\sqrt{2}$ in deuterium water.
The 3200–3400 cm.$^{-1}$ band is ν_3 of figure 2, since the frequency is only slightly less than the corresponding O—H frequency in water vapor. The 176 cm.$^{-1}$ band has been attributed to ν_1 and the 500 cm.$^{-1}$ band to ν_2.

The polarizability, α, of a harmonic vibration of charge e, mass m, and frequency ν is given by the expression

$$\alpha = \frac{e^2}{4\pi^2 m\nu^2} \tag{9}$$

The molal polarization, P, is

$$P = N\tfrac{4}{3}\pi\alpha = \frac{Ne^2}{3\pi m\nu^2} \tag{10}$$

From equation 10 the polarization per mole of hydrogen for the 500 cm.$^{-1}$ frequency is 39 cc. The number of hydrogen bonds in water is indefinite; hence the number of proton vibrators in a given direction in water is not definitely known but the maximum polarization per mole, taking into account the two hydrogens and the two-dimensional vibration, would be of the order of magnitude $2 \times \frac{2}{3} \times 39$ or 52. Even if only one-third of the possible hydrogen bonds are formed in water, this would give a polarization of 17 cc. From the index of refraction and the Clausius–Mosotti equation the electron polarization is 3.7 cc. It is obvious that this equation cannot be used to calculate the dielectric constant for high values of the proton polarization, but a polarization of 17 cc. appears adequate to account for the value of ϵ_0 observed by Collie, Hasted, and Ritson. The polarization for a deuterium vibration would be the same, since the decrease in the frequency by the factor of the square root of 2 balances the increase in mass.

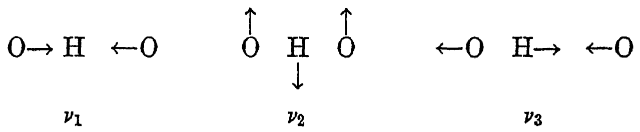


FIG. 2

As pointed out in the discussion of the Onsager equation, even a small value for the atomic polarization can thus produce a large change in ϵ_1, since it affects the proportionality constant between ϵ and P as indicated in equation 4. Indeed, the experimental value of ϵ_0 (5.5), obtained by Collie, Hasted, and Ritson for water, appears to be too high, as it would lead to a dielectric constant of at least 125. The contribution of the proton polarization will vary considerably from one hydrogen-bonded compound to another, depending upon the value for the vibrational frequency. In some compounds the frequency may be so high that the effect will be quite small. This may be true, for example, in KHF_2.

Hydrogen-bond formation also lengthens the X—H bond distance. From the change in Raman frequency, Sutherland (14) has calculated that the increase may be as large as 0.1 Å. for the O—H distance on association of two carboxyl groups. The dipole moment of a molecule such as water would then be appreciably increased over the value for the vapor, but in terms of the

Onsager equation, this effect on the dielectric constant of hydrogen-bonded liquid should not be as significant as the contribution from proton polarizations.

THE MECHANISM OF DIPOLE ORIENTATION IN ICE

An interpretation of the dielectric constant of ice must account for the following facts: The constant has a value of 75 at 273°C., and the relaxation time is between 10^{-5} and 10^{-6} sec. at that temperature but rapidly increases at lower temperatures. For comparison the relaxation time in water is around 10^{-10} sec. For frequencies greater than 10 kilocycles the dielectric constant falls to a small value (less than 10) at low temperatures, but there is evidence that the constant for infinite wave length does not decrease. Thus Murphy (10) charged an ice condenser for 1 hr. at −138.7°C. and by successive discharges collected about the same total charge as is obtained at 0°C. Also by rapidly cooling an ice condenser with liquid air, it is possible to freeze a permanent electric moment into the ice, which again indicates that a dipole must go over an appreciable

TABLE 4

ΔE and ΔS for the dipole orientation activation in ice

EXPERIMENTER	ΔE	ΔS
	kcal./mole	*cal./degree/mole*
Wentsch	9.3	0
Murphy	14.4	17
Smyth and Hitchcock	12.2	13

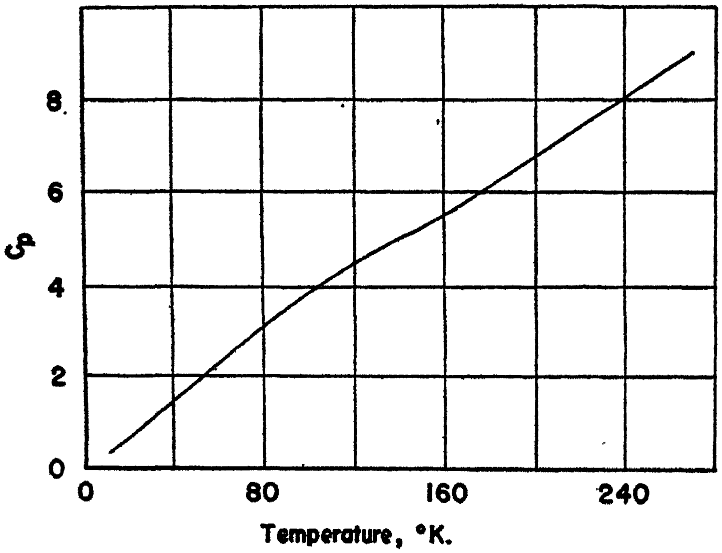


FIG. 3. Heat capacity of ice; C_p in calories per degree per mole

barrier in returning to its normal position. Quantitative information on the activation energy of the orientation barrier can be obtained from the temperature coefficient of the relaxation time. Kauzmann (6; see also 13) has calculated the values summarized in table 4. The specific heat curve obtained by

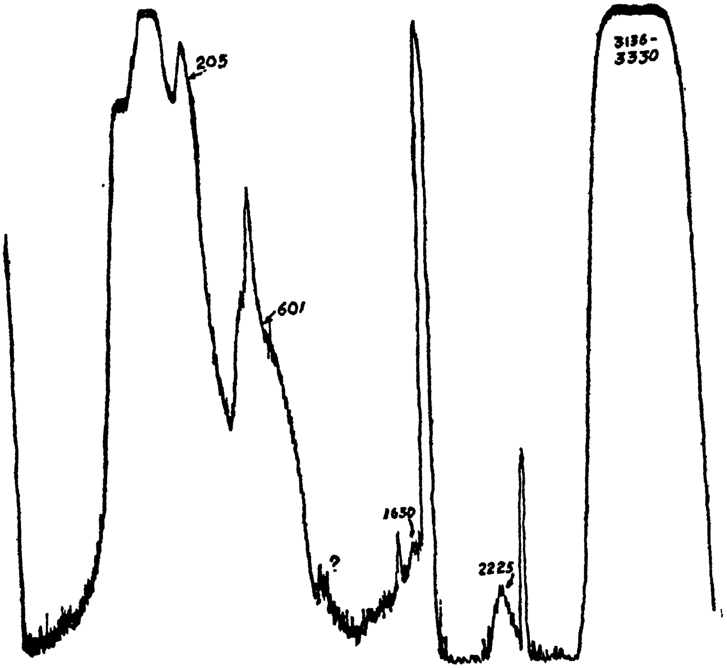


FIG. 4. Raman spectrum of ice (from Hibben)

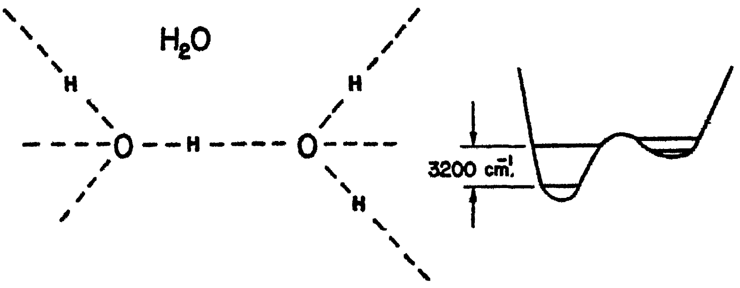


FIG. 5. Potential diagram for the proton in the hydrogen bond in ice

Giauque and Stout (4) is given in figure 3. Around 120°K. the curve looks as though it would extrapolate to a value of $C_p = 6$, which would correspond to the equipartition value for the oxygen, and then it rises in almost a straight line to the value of 9, or 1.5 cal. per hydrogen higher. This rise is not the sort of curve one expects with the acquisition of rotational energy.

The Raman spectrum, again from Hibben, is given in figure 4. It has the same general characteristic as the water spectrum but the 500 cm.$^{-1}$ bond is located at 600 cm.$^{-1}$ and is sharper. Presumably it is due to the hydrogen vibrations at right angles to the bonds. At 273°K. such a two-dimensional vibration with $\omega/T = 2.2$ would have a specific heat of 1.8 cal. per gram-atom of hydrogen, which is approximately the excess over 6 noted above.

If the water molecules in ice are not rotating, the question arises as to how it is possible to obtain the dipole orientation which is required for the high dielectric constant. In answer to this question the following mechanism is suggested. The frequency of the first vibrational state, ν_3 of figure 2, is 3200 cm.$^{-1}$ from the

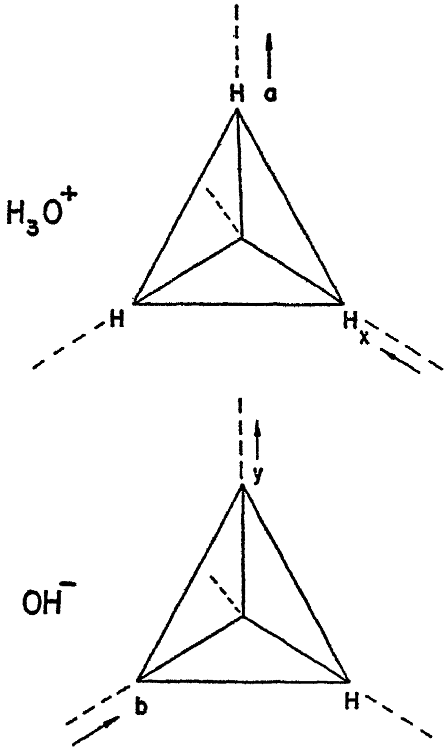


Fig. 6. Mechanism for the dipole orientation

Raman spectrum. This corresponds to 9100 cal./mole. Protons in this state should be able to cross the barrier, as is indicated in figure 5. While the population of molecules in the resulting H_3O^+ and OH^- states will be quite small, the rate at which such ion pairs are forming and disappearing will be quite rapid. This process provides a mechanism for dipole reorientation, since in the H_3O^+ ion there is the possibility of any one of the three protons jumping off and on the OH^- the possibility of any one of three protons returning to form the neutral molecule. As illustrated in figure 6, if the H_2O gains a proton at X and loses one at a, the dipole is reoriented. Likewise a loss at y and a gain at b produce a similar orientation. Ordinarily the same proton which jumped over will jump back, but there certainly will be conditions on the neighboring molecules which will permit the optional process and under the influence of an electric field a certain amount of orientation in the direction of the field will occur. The value of 9100 cal. for the vibrational excitation agrees with the experimental value of Wentsch (9300 cal., table 4) and such a process would have a very small entropy change. It would be difficult to account for an entropy of activation as high as 17 cal./degree, which was calculated from the work of Murphy, since the total entropy of an ice molecule at 273°K. is only 9.1 cal./degree.

SUMMARY

In 1920 Latimer and Rodebush (8), in their paper in which they introduced the idea of the hydrogen bond, made the suggestion that the displacement of the proton in the bond was doubtless related to the high dielectric constant. This relationship may now be summarized.

1. An appreciable displacement of the proton along the direction of the bond increases the permanent moment of the molecule.
2. A small atomic polarization of the proton in an electric field results in a big change in the dielectric constant, as indicated by the Onsager equation.
3. A transfer of the proton over the barrier offers a mechanism for the orientation of the dipoles by an electric field in ice and probably other solids.

REFERENCES

(1) Collie, C. H., Hasted, J. B., and Ritson, P. M.: Proc. Phys. Soc. (London) 73 (1948).
(2) Coop, I. E., Davidson, N. R., and Sutton, L. E.: J. Chem. Phys. **6**, 905 (1938).
(3) Debye, P.: *Polar Molecules.* The Chemical Catalog Company, Inc., New York (1929).
(4) Giauque, W. F., and Stout, J. W.: J. Am. Chem. Soc. **58**, 1144 (1936).
(5) Hibben, J. H.: *The Raman Effect.* The Reinhold Publishing Corporation, New York (1939).
(6) Kauzmann, W.: Rev. Modern Phys. **14**, 12 (1942).
(7) Kirkwood, J. G.: J. Chem. Phys. **7**, 911 (1939).
(8) Latimer, W. M., and Rodebush, W. H.: J. Am. Chem. Soc. **42**, 1419 (1920).
(9) Maryott, A. A., Hobbs, M. E., and Gross, P. M.: J. Chem. Phys. **8**, 415 (1946).
(10) Murphy, E. J.: Trans. Am. Electrochem. Soc. **65**, 133 (1931).
(11) Onsager, L.: J. Am. Chem. Soc. **58**, 1486 (1936).
(12) Pohl, H. A., Hobbs, M. E., and Gross, P. M.: J. Chem. Phys. **8**, 424 (1946).
(13) Powell, R. E., and Eyring, H.: Advances in Colloid Sci. **1**, 183 (1942).
(14) Sutherland, G. B. B. M.: Proc. Indian Acad. Sci. **8**, 341 (1938).
(15) Wyman, J.: J. Am. Chem. Soc. **58**, 1482 (1936).

CRITICAL PHENOMENA IN BINARY LIQUID SYSTEMS[1]

O. K. RICE

Department of Chemistry, University of North Carolina, Chapel Hill, North Carolina

Received October 25, 1948

CONTENTS

I. Critical phenomena in one-component systems 69
II. Formulation of the problem for a two-component system........ 71
III. General thermodynamic considerations........ 72
IV. Association theory for the binary system........ 76
  General theory........ 76
  Critical point with horizontal portion of coexistence curve........ 79
  One-phase region outside a horizontal coexistence curve........ 82
  Critical point with a rounded coexistence curve........ 83
V. The experimental material........ 84
  General observations........ 84
  The shape of the coexistence curve........ 85
  The chemical potential in the critical region........ 87
  Effect of addition of a third component........ 89
  Suggestions for future experiments........ 90
VI. Summary........ 90
VII. References........ 92

The phase relationships in a two-component liquid system are generally expressed in a temperature–composition diagram for constant (usually essentially zero) pressure, by a coexistence curve which gives the composition of the two liquid phases in equilibrium with each other at any temperature. Frequently there is a temperature above which (upper critical solution temperature) or below which (lower critical solution temperature) the two liquids are completely miscible, forming only one phase regardless of composition. It will be the object of this paper to consider the phenomena which occur in the neighborhood of such a critical point. In order to understand the problems involved, it will be well to start with a brief review of critical phenomena in the vapor–liquid equilibrium of one-component systems, which have many points of similarity.

I. CRITICAL PHENOMENA IN ONE-COMPONENT SYSTEMS

In a one-component system, the coexistence curve gives the molal volume (inverse of concentration) of the liquid and vapor phases which are in equilibrium at any temperature or any pressure. According to the van der Waals theory, the coexistence curve has a rounded, roughly parabolic shape in the neighborhood of the critical point. However, careful observations of the meniscus separating vapor and liquid phases in a number of cases (21, 26, 35; see the last of these for earlier references) indicate that in a diagram in which either pressure or temperature is plotted as ordinate against molal volume as abscissa, the coexistence curve has a flat top, having the appearance of a truncated parabola. There is a range of molal volumes (volume of 1 mole of the entire two-phase system) over which the meniscus disappears, within the limits of error, at exactly the same temperature, which is designated as T_m.

---

[1] Presented at the Symposium on Thermodynamics and Molecular Structure of Solutions, which was held under the auspices of the Division of Physical and Inorganic Chemistry at the 114th Meeting of the American Chemical Society, Portland, Oregon, September 13 and 14, 1948.

This observation is in accord with the results obtained from the statistical-mechanical theory of condensation of Mayer and Harrison (11, 22). They also believed, on the basis of this theory, that at temperatures just above T_m the isotherms should have a horizontal (constant-pressure) portion extending over a range of volumes.

McIntosh, Dacey, and Maass (5, 19) believed that they observed such horizontal portions in the isotherms of ethylene above the temperature at which the meniscus could be observed, but their data can be fitted just as well by curves with finite slopes in this region (31). Data on carbon dioxide (23) definitely seem to indicate that the horizontal portion is of finite length in the last isotherm which has a horizontal portion, and that only 0.15° above the critical temperature the isotherms have everywhere a finite slope. However, in this case the meniscus was not observed, so the question cannot be considered to be entirely settled experimentally.

Recently (31) I have considered the process of condensation of a vapor from the point of view of associating molecules (2, 8, 30). This involves the consideration of the equilibrium between single molecules, double molecules, triples, and higher clusters or droplets, the molecules within a cluster being held together by van der Waals or, possibly, dipole forces. As the system is compressed to smaller volumes, more and more of the larger clusters are formed. If it is compressed sufficiently clusters of macroscopic size suddenly become stable, i.e., condensation begins.

However, events will proceed in this way only if the surface tension of the droplets is positive. The surface tension of a liquid decreases with increasing temperature and will eventually approach zero. Then the sudden shift of stability to the large clusters will not occur. Instead, the number of large clusters and their average size will gradually increase until finally, when the *free* volume of the large clusters is reduced to zero, they coalesce into one. The free volume of a large cluster is determined, not by the freedom of motion of the individual molecules within it, but by the limits within which the cluster as a whole can move without becoming joined to another cluster. Thus the free volume of the large clusters may be reduced to zero when the density of the system is still rather low. The coalescence of the droplets does not result in a phase which is separated by a meniscus from the vapor, but, instead, it will produce a spongy, complexly interconnected mass of variable density. The process of condensation will now consist of the compression of the spongy mass, reducing the at first interconnected regions of low density to the status of disconnected bubbles in the liquid. In this process there will be a decrease in the surface, but there will be no resistance to the compression, since the surface tension is zero, and the pressure will remain constant until this process is finished; further compression requires actual squeezing of the liquid, rather than squeezing out of "gaseous" portions of the mass. The process which takes place at constant pressure presumably corresponds to the change of volume across the flat top of the coexistence curve.

Above the temperature T_m the surface tension will become negative. Once the spongy mass is formed, it will require a small but steady increase in pressure in order to squeeze out the interstices, for this requires a decrease in the total amount of surface, which is now resisted on account of the negative surface tension. Thus, according to this view, there will be no constant-pressure portion of an isotherm above the temperature T_m. This apparent contradiction to the statistical theory of Mayer and Harrison (11, 22) may, in my opinion, be explained in the following way (31). In the theory of Mayer and Harrison there appear certain cluster integrals, which are assumed to be volume independent. The integrals which involve numbers of molecules comparable to the total number will, however, become volume dependent when the volume actually filled by these molecules is comparable to the total volume. If these integrals are then important in Mayer and Harrison's equation, their theory will break down. It is difficult to prove at just what point this breakdown occurs, but it is reasonable to believe that it happens when the spongy mass described above is formed. If this is the case, it is shown in the previously cited work, that this will—or at least may—cause Mayer and Harrison's theory to break down before the flat portion of the isotherm, predicted in their theory, is reached.

II. FORMULATION OF THE PROBLEM FOR A TWO-COMPONENT SYSTEM

In many ways the two-phase two-component system of two partially immiscible liquids is analogous to the two-phase one-component system. With the total pressure of the two-component system held constant, the number of degrees of freedom is the same. The concentration of one of the components takes the place of the molal volume (actually it is analogous to the inverse of volume) of the one-component system, and the fugacity of this component is analogous to the pressure of the one-component system. (Often it is more convenient to work directly with the chemical potential.)

We can therefore see at once the questions which arise in the case of the two-component system. These are twofold: (*1*) Does the coexistence curve have a horizontal portion (constant temperature over a range of concentrations) at the critical temperature? (*2*) Just above or below the critical temperature, in the region where there is apparently complete miscibility, is there a range of temperatures for which an isotherm shows a constant fugacity or chemical potential for one component (and hence for the other also) over a range of concentrations? It will be the purpose of this article to consider these questions from both the theoretical and the experimental points of view.

III. GENERAL THERMODYNAMIC CONSIDERATIONS

We shall set up the following notation in dealing with a two-phase binary liquid system (see figure 1):

$'$ refers to one of the phases, "phase-prime"
$''$ refers to the other phase, "phase-double-prime"
Subscript a refers to component A
Subscript b refers to component B
x_a', x_b', x_a'', x_b'' are mole fractions
By definition $x_a' \geq x_a''$ and $x_b'' \geq x_b'$, so
  Component A is solvent for phase-prime and
  Component B is solvent for phase-double-prime
$\mu_a', \mu_b', \mu_a'', \mu_b''$ are chemical potentials
A' is $\dfrac{1}{RT}\left(\dfrac{\partial\mu_a'}{\partial \ln x_a'}\right)_{T,p} = \dfrac{1}{RT}\left(\dfrac{\partial\mu_b'}{\partial \ln x_b'}\right)_{T,p}$
A'' is $\dfrac{1}{RT}\left(\dfrac{\partial\mu_a''}{\partial \ln x_a''}\right)_{T,p} = \dfrac{1}{RT}\left(\dfrac{\partial\mu_b''}{\partial \ln x_b''}\right)_{T,p}$
  (The relations follow from the Gibbs–Duhem equation)
$\bar{s}_a', \bar{s}_b', \bar{s}_a'', \bar{s}_b''$ are partial molal entropies
$\Delta S_a = \bar{s}_a'' - \bar{s}_a'$ and $\Delta S_b = \bar{s}_b' - \bar{s}_b''$

From the general equation for constant pressure, written with either subscript and either prime

$$d\mu = -\bar{s}\, dT + RTA\, d \ln x \tag{1}$$

and from the general set of conditions of equilibrium

$$\mu_a' = \mu_a''; \qquad \mu_b' = \mu_b'' \tag{2}$$
$$d\mu_a' = d\mu_a''; \qquad d\mu_b' = d\mu_b''$$

we obtain (18):

$$0 = -\Delta S_a dT - A'RT\, d \ln x_a' + A''RT\, d \ln x_a'' \tag{3}$$
$$0 = \Delta S_b dT - A'RT\, d \ln x_b' + A''RT\, d \ln x_b''$$

In these equations we have used the total differentiation sign to mean differentiation along the coexistence curve. This convention will be used throughout the argument leading up to equation 10. In any differential of this sort it is understood that the displacement along one branch of the coexistence curve corresponds to (implies the same dT) as displacement along the other branch.

Since $dx_a' = -dx_b'$ and $dx_a'' = -dx_b''$, we can write

$$d \ln x_a'' = -(x_b''/x_a'')\, d \ln x_b'' \tag{4}$$
$$d \ln x_b' = -(x_a'/x_b')\, d \ln x_a'$$

Using these in equation 3 and solving the simultaneous equations, we obtain

$$\frac{\mathrm{d} \ln x'_a}{\mathrm{d}T} = \frac{\Delta S_a - (x''_b/x''_a)\Delta S_b}{A'RT(x'_a x''_b/x'_b x''_a - 1)} \tag{5}$$

and

$$\frac{\mathrm{d} \ln x''_b}{\mathrm{d}T} = \frac{\Delta S_b - (x'_a/x'_b)\Delta S_a}{A''RT(x'_a x''_b/x'_b x''_a - 1)} \tag{6}$$

These equations determine the coexistence curve. We shall now consider the necessary conditions that the coexistence curve should have an upper, or a lower, critical temperature. To fix the ideas, let us first consider the case of an upper critical temperature, and let us suppose that the curve does *not* have a flat top but is curved in the region of the critical temperature. Then both d ln $x'_a/\mathrm{d}T$ and d ln $x''_b/\mathrm{d}T$ will be infinite at the critical point and they will be negative in the immediate neighborhood of the critical point, as will be clear from figure 1.

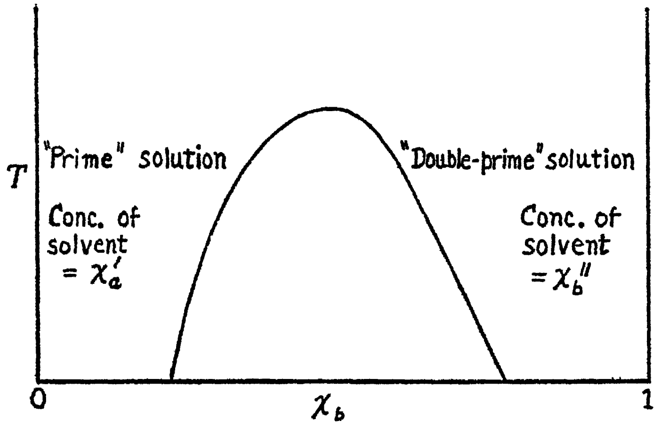


FIG. 1. Schematic diagram of a coexistence curve, illustrating the notation used

Let us now consider the behavior of the quantities in the fractions on the right-hand sides of equations 5 and 6. A' and A'' cannot be negative in any physically realizable state of the system. Since $x'_a \geq x''_a$ and $x''_b \geq x'_b$, the factor $x'_a x''_b/x'_b x''_a - 1$ must be greater than or equal to zero. Thus the denominators must be positive or zero; hence the numerators must be negative or zero in the neighborhood of an upper critical point. At the critical point where $x'_a = x''_a$ and $x'_b = x''_b$ they clearly must be zero. We also see that

$$\frac{\mathrm{d}}{\mathrm{d}(x''_b/x''_a)}\left(\frac{x'_a x''_b}{x'_b x''_a} - 1\right) = \frac{x'_a}{x'_b} + \frac{x''_b}{x''_a}\,\frac{\mathrm{d}(x'_a/x'_b)}{\mathrm{d}(x''_b/x''_a)}$$

$\mathrm{d}(x'_a/x'_b)/\mathrm{d}(x''_b/x''_a)$ must be positive; thus the above expression is positive and $x'_a x''_b/x'_b x''_a - 1$ vanishes to the first order at the critical point.

Let us now consider the numerator of equation 5, which we shall designate as

$\mathfrak{h}$. Differentiating along the coexistence curve, remembering the definitions of ΔS_a and ΔS_b, we have

$$\begin{aligned}\frac{d\mathfrak{h}}{d(x_b''/x_a'')} = \left(\frac{\partial \bar{s}_a''}{\partial(x_b''/x_a'')}\right)_T - \left(\frac{\partial \bar{s}_a'}{\partial(x_b'/x_a')}\right)_T \frac{d(x_b'/x_a')}{d(x_b''/x_a'')} \\ - \frac{x_b''}{x_a''}\left(\frac{\partial \bar{s}_b'}{\partial(x_b'/x_a')}\right)_T \frac{d(x_b'/x_a')}{d(x_b''/x_a'')} + \frac{x_b''}{x_a''}\left(\frac{\partial \bar{s}_b''}{\partial(x_b''/x_a'')}\right)_T - \Delta S_b \\ + \frac{\partial \Delta S_a}{\partial T}\frac{dT}{d(x_b''/x_a'')} - \frac{x_b''}{x_a''}\frac{\partial \Delta S_b}{\partial T}\frac{dT}{d(x_b''/x_a'')} \qquad (7)\end{aligned}$$

The partial derivatives with respect to T are taken at constant mole fraction. Since, at the critical point, $x_a' = x_a''$ and $x_b' = x_b''$, both $\partial\Delta S_a/\partial T$ and $\partial\Delta S_b/\partial T$ are zero. The factor $dT/d(x_b''/x_a'')$ is also zero. Hence the last two terms in equation 7 will not even contribute to the second derivative at the critical point and may be completely neglected. At constant temperature we have the general relation

$$x_a'' d\bar{s}_a'' + x_b'' d\bar{s}_b'' = x_a' d\bar{s}_a' + x_b' d\bar{s}_b' = 0$$

Hence equation 7 reduces to

$$\frac{d\mathfrak{h}}{d(x_b''/x_a'')} = \left(\frac{x_b'}{x_a'} - \frac{x_b''}{x_a''}\right)\left(\frac{\partial \bar{s}_b'}{\partial(x_b'/x_a')}\right)_T \frac{d(x_b'/x_a')}{d(x_b''/x_a'')} - \Delta S_b \qquad (8)$$

At the critical point $x_b'/x_a' - x_b''/x_a''$ and ΔS_b are both zero. Hence this expression vanishes.

Omitting terms which vanish at the critical point, the second derivative is

$$\frac{d^2\mathfrak{h}}{d(x_b''/x_a'')^2} = \left(\frac{d(x_b'/x_a')}{d(x_b''/x_a'')} - 1\right)\left(\frac{\partial \bar{s}_b'}{\partial(x_b'/x_a')}\right)_T \frac{d(x_b'/x_a')}{d(x_b''/x_a'')} - \left(\frac{\partial \Delta S_b}{\partial(x_b''/x_a'')}\right)_T \qquad (9)$$

On account of the symmetry of the curve in the immediate neighborhood of the critical point $d(x_b''/x_a'') = -d(x_b'/x_a')$. Hence equation 9 becomes

$$\frac{d^2\mathfrak{h}}{d(x_b''/x_a'')^2} = -2\left(\frac{\partial \bar{s}_b'}{\partial(x_b''/x_a'')}\right)_T - \left(\frac{\partial \Delta S_b}{\partial(x_b''/x_a'')}\right)_T \qquad (9a)$$

Actually S_b'' and S_b' are the same quantity; one of them is merely evaluated on one side of the coexistence curve, the other on the other. In the same sense, x_b''/x_a'' and x_b'/x_a' are essentially the same variable. Therefore at the critical point

$$[\partial\bar{s}_b'/\partial(x_b'/x_a')]_T \equiv [\partial\bar{s}_b''/\partial(x_b''/x_a'')]_T$$

But since along the coexistence curve $d(x_b'/dx_a')/d(x_b''/x_a'') = -1$, we see that

$$[\partial\bar{s}_b'/\partial(x_b''/x_a'')]_T = [\partial\bar{s}_b'/\partial(x_b'/x_a')]_T\, d(x_b'/x_a')/d(x_b''/x_a'') = -[\partial\bar{s}_b''/\partial(x_b''/x_a'')]_T$$

With this we can reduce equation 9a to

$$\frac{d^2\mathfrak{h}}{d(x_b''/x_a'')^2} = -2\left(\frac{\partial \Delta S_b}{\partial(x_b''/x_a'')}\right)_T \qquad (10)$$

Now we have seen that $\mathfrak{h} = 0$ at the critical point and must be negative in the neighborhood of an upper critical solution point. Since $d\mathfrak{h}/d(x_b''/x_a'')$ is also zero at the critical point the second derivative must be negative. Since x_b''/x_a'' increases on leaving the critical point along the coexistence curve, it is seen that this requires that ΔS_b increase also. Since ΔS_b starts at zero, it must be positive near an upper critical solution point. A similar discussion of equation 6 will show that ΔS_a must also be positive near an upper critical point, and both ΔS_a and ΔS_b will be negative near a lower critical point.

Furthermore, since the first derivatives of the numerators vanish, the numerators themselves must vanish to the second order at the critical point. Since the factor $x_a'x_b''/x_b'x_a'' - 1$ in the denominator vanishes only to the first order, this means that A' and A'' must vanish to the second order in order for the left-hand side of equations 5 and 6 to become infinite to the first order. This corresponds to the vanishing of both $(\partial p/\partial V)_T$ and $(\partial^2 p/\partial V^2)_T$ at the critical point according to the van der Waals theory, in a one-component system.

The signs of the ΔS's can be interpreted in terms of the molecular structure of the liquids. If ΔS_a and ΔS_b are positive, this means an increase of entropy on transfer of material from a more concentrated to a less concentrated solution. This, of course, is the normal expectation, and results in an upper critical temperature. If ΔS_a and ΔS_b are negative this can result only from special interactions between the two kinds of molecules, such as hydrogen bonds, which are known to be instrumental in causing lower critical solution temperatures (13). If a molecule introduced as a solute into another liquid can cause a considerable degree of orientation of the solvent molecules, this orientation may lower the entropy of the solution to such an extent that transfer of the solute molecule from a concentrated to a dilute solution can actually result in a decrease of entropy. Since the interaction between the two kinds of molecules is mutual, it is not at all surprising that ΔS_a and ΔS_b should be negative at the same time. If they had opposite signs the coexistence curve would have neither an upper nor a lower critical temperature, and this is also a possibility which must be reckoned with.

Some pairs of liquids show both an upper and a lower critical point. It would be naturally expected that the orientation effect would be most marked at a low temperature, so it is possible for ΔS_a and ΔS_b to change sign in such a way that there is a lower critical temperature and a higher upper critical temperature, the coexistence curve then being a closed curve.

The above discussion deals only with the case where the coexistence curve is rounded at the critical temperature. If the coexistence curve is flat at the critical temperature, $\mathrm{d}\ln x_a'/\mathrm{d}T$ and $\mathrm{d}\ln x_b''/\mathrm{d}T$ are infinite for a range of concentrations, so that we do not in general have $x_a' = x_a''$ and $x_b' = x_b''$ at the critical temperature. Then the numerators of equations 5 and 6 will be finite and the factor $x_a'x_b''/x_b'x_a'' - 1$ will be finite and positive. It is thus not necessary for A' and A'' to vanish to the second order at the ends of the range of concentrations over which the coexistence curve is flat. If the two branches of the coexistence curve are to turn toward each other (as they always do) as the critical temperature is approached, then for an upper critical temperature we would have

$$\Delta S_a - (x_b''/x_a'')\Delta S_b < 0$$

and

$$\Delta S_b - (x_a'/x_b')\Delta S_a < 0$$

Multiplying the second of these by x_b''/x_a'' and adding to the first we obtain

$$(1 - x_b''x_a'/x_a''x_b')\Delta S_a < 0$$

Since the term in parentheses is always less than zero, we find $\Delta S_a > 0$. Similarly $\Delta S_b > 0$. These then represent *necessary* conditions for coexistence curves of normal form at an upper critical temperature, regardless of whether the top of the curve is flat or rounded. Similarly $\Delta S_a < 0$ and $\Delta S_b < 0$ are in general necessary conditions for a lower critical temperature.

IV. ASSOCIATION THEORY FOR THE BINARY SYSTEM

McMillan and Mayer (20) have developed a statistical theory of multicomponent systems which can be applied to binary liquid systems. Their conclusions are very similar to those of Mayer and Harrison (11, 22) on the vapor–liquid critical point. The vanishing of A' to the second order involves two independent conditions, which McMillan and Mayer believe could be satisfied simultaneously only if the attractive forces were of a very special type. However, they apparently hesitate to exclude entirely the possibility of a critical point with a rounded coexistence curve. The normal type of coexistence curve would seem, nevertheless, to be one with a horizontal section of the critical temperature. In this case, according to McMillan and Mayer, the chemical potentials should be constant over a finite range of compositions even beyond the critical point, that is, at temperatures at which the two liquids form a single homogeneous phase. However, in this region the theory will meet with the same difficulty as does the theory of Mayer and Harrison above the critical point of a one-component system.

In this paper we shall adapt the theory of associating molecules to a binary system. This theory is not as rigorous as that of McMillan and Mayer in the region where two phases are formed; I believe, however, that it is better for the region where only one phase is formed. In any case, the mathematical formulation is much simpler. Since the theory has already been developed in detail for one component systems (31) we can make some of the discussions brief, referring the reader to the previously published work for details.

General theory

The theory will be the same for either the prime or the double-prime phase as described in Section III. Let us consider the former, in which component B is the solute and component A the solvent. Component B can occur as single molecules, as pairs, as triplets, etc., in general as clusters of s molecules each. Within these clusters of species B there may exist in turn molecules or clusters of species A. However, a cluster will be designated by the number of molecules of species B. We let the number of single molecules of B in a volume V be designated as n_{b1}, the number of pairs as n_{b2}, $\cdots$, the number of clusters of s molecules of species B as $n_{bs}, \cdots$. Then

$$\Sigma_s s n_{bs} = N_b \tag{11}$$

where N_b is the total number of B molecules. We let the concentrations (n_{b1}/V, $n_{b2}/V, \cdots n_{bs}/V, \cdots$) be designated as $c_{b1}, c_{b2} \cdots c_{bs}, \cdots$, and the respective chemical potentials (referring to single, rather than moles of, clusters) be denoted by $\mu_{b1}, \mu_{b2}, \cdots, \mu_{bs}, \cdots$. Then if the solution is sufficiently dilute so that the laws of dilute solutions may be considered valid,

$$\mu_{bs} = s\phi_{bs} + kT \ln c_{bs} \tag{12}$$

where $s\phi_{bs}$ is the value of μ_{bs} when $c_{bs} = 1$. We now define a quantity β_b by the equation

$$\beta_b = \phi_{b\infty} - \phi_{b1} \tag{13}$$

β_b, then, is the free energy of "condensation" per molecule, of single molecules at unit concentration to form a very large cluster.

If we form a cluster of s molecules of B from single molecules, this may be considered as the condensation of $s - 1$ molecules on one molecule, and the free energy of condensation at unit concentration would be $\beta_b(s - 1)$, except for some corrections. In the first place, the free energy of condensation of the molecules added to a small cluster is not typical of those added later if the cluster is made to grow. This difference in free energy will persist in the total energy even for a large cluster.[2] This correction we shall call ϵ_{bs}. For large s it should approach a limit which we shall call $\epsilon_{b\infty}$. There will also be a correction for surface, or, rather, interfacial energy of the form $\gamma s^{2/3}$, where γ is a constant related to, but not exactly equal to, the interfacial tension. This form will also break down for small s, but we may take any correction into ϵ_{bs}. Collecting all these terms we get for the free energy of condensation of s single molecules into a single cluster, both types being at unit concentration,

$$s\phi_{bs} - s\phi_{b1} = \beta_b(s - 1) + \gamma s^{2/3} + \epsilon_{bs} \tag{14}$$

For the special case $s = 1$, we see that equation 14 gives:

$$\epsilon_{b1} = -\gamma \tag{15}$$

The condition of equilibrium is that the total free energy F of the system should be a minimum. Hence

$$\delta F = \Sigma_s \mu_{bs} \delta n_{bs} = 0 \tag{16}$$

subject to the condition, which follows from equation 11, that

$$\Sigma_s s \delta n_{bs} = 0 \tag{17}$$

In equation 16 we have no term which explicitly refers to the solvent. In view of the fact that a droplet or cluster of solute molecules may have in turn solvent molecules dissolved in it, this may seem to be an oversight. However, the solvent within any given cluster is assumed always to be in equilibrium with its surroundings. Therefore it is only necessary to impose the condition 16 on the number of solute clusters of various sizes. The effect of solvent is taken care of implicitly in the expression for $s\phi_{bs}$ to be obtained from equation 14.

---

[2] In the work on condensation of a vapor (31) I assumed ϵ_∞ to be zero. It is seen that this is not correct, but actually it does not affect the conclusions of my earlier paper. This will be discussed elsewhere.

From equations 16 and 17 we obtain by the standard procedure

$$\mu_{bs} + \alpha_b' s = 0 \tag{18}$$

where α_b' is a constant to be determined (the Lagrange multiplier). With equations 12 and 14 this gives

$$c_{bs} = \exp[(\beta_b - \alpha_b s - \gamma s^{2/3} - \epsilon_{bs})/kT] \tag{19}$$

where

$$\alpha_b = \phi_{b1} + \beta_b + \alpha_b'$$

In particular, remembering equation 15,

$$c_{b1} = e^{(\beta_b - \alpha_b)/kT} \tag{20}$$

α_b is determined by equation 11, or it may be considered to be determined by equation 20, all the c_{bs} being determined in terms of c_{b1}. α_b may be interpreted as the difference between the chemical potential $\phi_{b\infty}$ of very large drops and that of the prime solution as it actually exists. We may build up the solute in the prime solution by small additions of single unassociated molecules of component B. Between the small additions we allow equilibrium between clusters to be reëstablished; this takes place with no free energy change, since we never depart appreciably from equilibrium. Hence

$$\mu_b' = \mu_{b1} = \phi_{b1} + kT \ln c_{b1}$$

$$= \phi_{b1} + \beta_b - \alpha_b = \phi_{b\infty} - \alpha_b$$

Hence, as was to be shown,

$$\alpha_b = \phi_{b\infty} - \mu_b' \tag{21}$$

$\phi_{b\infty}$ may be taken as the chemical potential per molecule of B of the double-prime phase at the concentration given by the coexistence curve at temperature T. An infinitely large (macroscopic) drop of component B contains within itself droplets of component A. If the concentration of A within this large drop of B exceeds a certain value (as it would tend to if the over-all concentration of the solution were on the prime side of the coexistence curve), component A will condense out as a macroscopic droplet within the drop of B. The free energy of transfer of single B molecules at unit concentration to form a macroscopic cluster will, therefore, be the same as the free energy of transfer into the double-prime phase; in other words $\phi_{b\infty} = \mu_b''$. The concentration term in equation 12 is unimportant for $s = \infty$.

Critical point with horizontal portion of coexistence curve

One condition that the two macroscopic phases should be in equilibrium with each other is that $\phi_{b\infty} = \mu_b''$ should be equal to μ_b', hence $\alpha_b = 0$. Let us now examine equation 19 with this condition in mind. In the first place we note that a more positive value of α_b corresponds to a smaller concentration. As long as α_b is positive we see from equation 19 that large clusters are very improbable. Even when α_b is zero, the $\gamma s^{2/3}$ term makes the large clusters improbable, provided γ is positive. As soon as α_b becomes ever so slightly negative the expression 19 ceases to converge; that is to say, the large clusters suddenly become extremely probable. This corresponds to the formation of a new phase. The one-phase system is stable as long as α_b is positive.

The above deductions are based on the assumption that the ideal solution laws hold. It is clear that qualitatively the same results would be obtained if this were not the case. However, deviations from ideal solution laws become important when γ is close to zero. This has been discussed in detail in reference 31 and essentially the same method of handling the situation can be applied here. Equation 19 will still hold provided the concentrations are considered to be concentrations in free volume. This is not the same as replacing concentration by activity. Replacing concentration by concentration in free volume corrects only for the entropy effects; energy effects are implicitly taken care of by changes in the values of β_b and the ϵ_{b_s}. The free volume may, in a given solution, be different for clusters of different sizes, but we assume that it approaches a definite limit for large s. The characteristic thing about this limiting value of the free volume is that it can approach zero while the concentration is still finite. This occurs when the large clusters are compressed so closely together that they can no longer exist as separate entities but are forced to coalesce. It is to be noted that only the effect of the clusters on each other is to be considered. The volume occupied by the solvent molecules does not affect the free volume in this sense, for the concentrations at high dilution are defined in terms of the total volume of the solution.

We can now understand the situation when γ becomes zero. This case is of particular importance, since if γ becomes zero we shall no longer expect to be able to distinguish a meniscus, and we shall now show that the conditions for the occurrence of a critical temperature are fulfilled. This matter has been discussed in some detail, though from a more qualitative point of view, by Frenkel (7).

We define a total effective concentration in free volume as $C_b = N_b/V_f$, where V_f is the limiting value of the free volume for large clusters. If α is very small, so that most of the molecules of B are in large clusters, we may write

$$C_b = \Sigma_s s c_{bs} \tag{22}$$

If $\gamma = 0$, equation 19 becomes for large values of s

$$c_{bs} = e^{(\beta_b - \alpha_b s - \epsilon_{b\infty})/kT} \tag{23}$$

If we insert equation 23 into equation 22 we obtain an approximation, the sum in which is readily evaluable. With α_b/kT very small it takes the form

$$C_b = e^{(\beta_b-\epsilon_{b\infty})/kT}(kT/\alpha_b)^2 \tag{24}$$

which, solved for α_b, gives[3]

$$\alpha_b = kT[e^{(\beta_b-\epsilon_{b\infty})/kT}/C_b]^{1/2} \tag{25}$$

As we have noted, C_b is defined as N_b/V_f. Let us now investigate the properties of V_f. Let us suppose that N_b is constant and that the concentration of the solution is varied by changing the total volume, using a semipermeable piston through which solvent A can pass, but which is capable of controlling the volume V available to solute B. The free volume is the volume through which the center of a cluster is free to move, without coalescing with another cluster, and it may be considered as the cube of a free length. The latter may be set roughly equal, or at least proportional, to $V^{1/3} - V'^{1/3}$, where V' is the volume which would be occupied if the system were compressed, without change in the density of or the total amount of material in the large clusters, until the free volume just vanished. We thus write

$$\begin{aligned} V_f^{1/3} &= V^{1/3} - V'^{1/3} \\ &\cong (V - V')/3V'^{2/3} \end{aligned} \tag{26}$$

if $V - V'$ is small compared to V.

Actually, compression of the system may well cause a change in density of a cluster. And since the clusters are certainly not spherical when $\gamma = 0$ or even when γ is positive but very small, compression might change the effective volume occupied by a cluster by changing its shape. Thus V' will vary with V and, on account of the change in shape mentioned in the last sentence, we might even have $\mathrm{d}V'/\mathrm{d}V \sim 1$ in a highly concentrated solution. The actual value of V at which V_f vanishes, we shall call V_0'; at this volume $V_0' = V = V'$.

Equation 26 gives:

$$C_b = N_b/V_f = 27N_bV'^2/(V - V')^3 \tag{27}$$

Setting equation 27 into equation 25 we obtain:

$$\alpha_b = kTe^{(\beta_b-\epsilon_{b\infty})/2kT}(V - V')^{3/2}/(27N_b)^{1/2}V' \tag{28}$$

The terms $e^{(\beta_b-\epsilon_{b\infty})/2kT}$ and V' (in the denominator) will have a relatively small dependence on V as compared with $(V - V')^{3/2}$ when V is close to V'. We therefore can write approximately

$$(\partial\alpha_b/\partial V)_T = \tfrac{3}{2}kTe^{(\beta_b-\epsilon_{b\infty})/2kT}(27N_b)^{-1/2}(V - V')^{1/2}V'^{-1}\left(1 - \frac{\mathrm{d}V'}{\mathrm{d}V}\right) \tag{29}$$

We now see that $(\partial\alpha_b/\partial V)_T$ vanishes when $V = V' = V_0'$. By equation 21 this means that $(\partial\phi_{b\infty}/\partial V)_T - (\partial\mu_b'/\partial V)_T$ vanishes. In view of the discussion following equation 21, we conclude that $(\partial\phi_{b\infty}/\partial V)_T = 0$; hence at $V = V' = V_0'$ we also have $(\partial\mu_b'/\partial V)_T = 0$. Of course the argument applies just as well to the other side of the coexistence curve, where $(\partial\phi_{a\infty}/\partial V)_T$ and $(\partial\mu_a''/\partial V)_T$ will be equal to zero. This in turn means that A' and A'' of Section III will vanish. Thus it is seen that the conditions are fulfilled for a critical temperature, with the coexistence curve having a horizontal section,[3] since the left-hand sides of equations 5 and 6 will become infinite.

If the vanishing of $(\partial\alpha_b/\partial V)_T$ depends only on the term $(V - V')^{1/2}$ it vanishes to a low order. As indicated above it may vanish to a higher order on account of the term $1 - dV'/dV$. This has already been discussed for the liquid–vapor equilibrium of a one-phase system (in Section 5 of reference 31), and the arguments need not be repeated here. It may be added, however, that the above discussion depends on the assumption that we can always recognize the boundaries of a cluster, even when it is close to another cluster. There may actually be some haziness in this matter, which would lead to haziness in the concept and evaluation of the free volume. This might result in a further smoothing out and rounding off of the curve of chemical potential *vs.* concentration.

The manner in which the interfacial tension, or γ, which is proportional to it, vanishes is of some importance. Let us write

$$\gamma = \eta - T\sigma \tag{30}$$

where η and σ are, respectively, proportional to the interfacial energy (more exactly enthalpy) and entropy (the proportionality being the same as that which relates γ to the interfacial tension).

In general, when γ is positive η and σ have the same sign. If they are positive, then at a sufficiently high temperature γ vanishes, giving an upper critical temperature. If they are negative, then at a sufficiently *low* temperature γ vanishes, giving a lower critical temperature.

If there are no special forces involved between the molecules of components A and B, then we may in general expect σ to be positive. For the interface is then a place where the two species of molecules mix, and where there is more disorder than in the bodies of the solutions. Under the same circumstances we expect ΔS_1 and ΔS_2 to be positive, and from either σ or ΔS_1 and ΔS_2, we conclude that if we have any critical temperature it will be an *upper* critical temperature.

On the other hand, if there are special forces between the molecules there will be a tendency for the molecules of A and B to be bound to each other in a special way at the interface. This is likely to lead to a negative σ and a *lower* critical temperature, according to the criteria considered in Section III.

---

[3] In reference 31 we considered the possibility that the expression for the chemical potential should have an entropy term of the form $kT\ a \ln s$, where a is a constant which would remain even when γ was zero. Such a term would be of the nature of a surface tension, but one which varied with s and became smaller for larger s. Such a term would, of course, affect the form of equations 24 and 25 and could affect the conclusions to be drawn from these equations if a became of the order of magnitude of 1. Actually, of course, our whole development is based upon the assumption that there does exist a temperature at which any term which is not proportional to s becomes negligible for large s, or is counteracted by some other effect. It seems likely that the term $1 - dV'/dV$ in equation 29 has more effect on the shape of the μ_b *vs.* concentration curve than the effect considered here. These details of behavior at the critical temperature obviously require further study.

Of course, we cannot predict in any case whether or not a critical point will actually occur, but there is a general consistency between the two points of view.

One-phase region outside a horizontal coexistence curve

If the interfacial tension vanishes in the way we have just discussed, one important consequence is that it can continue in the same direction and become negative. Frenkel (7) has argued against this, but his arguments do not seem convincing to me. It is true that one cannot measure a negative interfacial tension, simply because one cannot find the interface. The interface tends to proliferate, so the clusters are no longer spherical, and in general we may expect the surface to become proportional to s rather than to $s^{2/3}$. This means that in equation 19 the surface term can be merged with the term $\alpha_b s$. Presumably, unless the relation of the surface quantities is such that sheets or filaments are stable,[4] proliferation of the surface will stop when the average curvature of the surface becomes sufficiently great. At this point the interfacial tension might be said to be zero in the sense that the free energy of the interface is a minimum with respect to changes in the surface which involve a change in average curvature. On the other hand, the free energy of the surface as a whole is negative, and the surface resists any attempt to lessen its area which leaves the average curvature unchanged. In any event, since proliferation of the interface is wasteful of free volume, it is probable that before the large clusters have coalesced into a spongy mass, the curvature of the surface will have become somewhat less than would be the case if the clusters were isolated.

The coalescence of the clusters of solute at the point where the free volume vanishes will produce a spongy mass of one phase in the other, just as in the case of liquid–vapor equilibrium there was produced a spongy mass of liquid in the vapor. We can treat the solution in much the same way as we treat the one-component system of condensing vapor, if we introduce a semipermeable piston, as above, which allows solvent A to pass through but holds back solute B. We assume osmotic equilibrium established between the molecules A outside and inside the piston. Let us now compress the piston slightly. A certain amount of A will flow out; since there is osmotic equilibrium this occurs without change of free energy. In order to maintain the concentration of phase rich in component A some material from this phase will enter into the spongy mass; as the system is in equilibrium this also occurs without free energy change. Finally the spongy mass will be compressed. This presumably lowers its surface area. If the interfacial tension is zero, there is no resistance to the piston on this account. If, however, the interfacial tension is negative there will be a resistance to this process. Equilibrium will be reëstablished if a higher pressure is applied to the piston. Suppose, however, we reëstablish the equilibrium by lowering the pressure of the solvent material A outside the piston, leaving the pressure inside the same as before. We then see that μ_a, the chemical potential of A, has decreased and hence μ_b will have increased. Thus if the interfacial tension is negative in the region beyond the critical point, we cannot expect μ_a or μ_b to be constant over any range of concentrations. Hence our theory predicts no range of concentrations in an apparently homogeneous liquid phase where the chemical potential will remain constant.

---

[4] If filaments and sheets should become stable, then since they constitute an entirely different state of the system, their appearance would manifest itself by a *sudden* drop of the interfacial tension to zero. This is, perhaps, a possibility which would be worthy of study, but we shall not consider it here. In general, however, when we go beyond the critical temperature into the range of homogeneous mixtures, we may expect the curvature of the interface to become greater and greater, so that sheets and filaments probably appear gradually and finally break up into single molecules, giving a microscopically homogeneous solution.

Critical point with a rounded coexistence curve

Thus far Section IV has dealt only with the case in which the coexistence curve has a horizontal stretch at the critical temperature. If the curve is, instead, rounded at the critical temperature, as in the case discussed in the first part of Section III, the compositions of the two phases finally become identical at the critical temperature. The interfacial tension between two identical phases would, of course, be zero, with both terms in equation 30 vanishing. However, two phases can have the same composition without being identical in every respect, if they contain clusters of varying composition, since only the average composition of such a microscopically heterogeneous phase would be measured. If, however, the clusters were of different composition than the body of the phase, the interfacial tension at the surface of the clusters would not in general be zero. But if this were true, large clusters would not form gradually; we could have a sudden appearance of a new phase, but the conditions for a normal critical point would not be fulfilled and we would be led to a contradiction. So if we do have a rounded coexistence curve the two phases do approach identity, even when the examination is made on a molecular scale. As they approach identity, and the interfacial tension becomes very small, clusters would of course form, but they would have almost the same composition as the rest of the liquid. At the critical point they would become "ghosts."

Suppose we consider a phase in which component B is solute and A solvent (the prime phase) at a temperature such that two phases can exist in equilibrium, just before the critical point is reached. If we dilute the solution by adding A, the clusters, which are almost ghosts, could not remain in equilibrium with the body of the liquid of the new composition. Since the clusters have almost the same composition as the surrounding medium and cannot become richer in A, a rather large amount of material will be withdrawn from them, to restore the original composition. Now a rounded coexistence curve requires that A' and A'', and hence $(\partial\alpha_b/\partial V)_T$ and $(\partial\alpha_a/\partial V)_T$, vanish to the second order, which, from equation 29, means that $V - V'$ should be proportional to $(V - V_0')^2$. This would require considerable swelling of the clusters to keep the volume nearly filled as V was increased by means of the semipermeable piston starting with $V = V' = V_0'$. That this swelling should be especially great just under the circumstance in which there is considerable loss of the material rich in B does not seem probable. On this basis we would not expect to find rounded coexistence curves. However, with the situation described it would certainly become very difficult to define the boundaries of a cluster, and with the identity of the clusters becoming increasingly hazy, it may well be that we can draw no valid conclusions on the basis of the cluster or association theory as developed here. McMillan and Mayer (20), as we have seen, concluded that the vanishing of A' and A'' to the second order would occur under only rather special circumstances, but it does not seem necessary to conclude that those circumstances are at all impossible, nor do the experimental data, discussed in the next section, seem adequate to decide this question unequivocally.

V. THE EXPERIMENTAL MATERIAL

The above sections have provided at least a partial theoretical answer to the questions propounded in Section II: (*1*) The coexistence curve will in many, probably most, cases have a horizontal section at the critical temperature, though this may not be absolutely necessary. (*2*) The chemical potentials should not be constant over any part of the region of complete miscibility.

General observations

It was long ago suggested (9, 27, 28, 29), on the basis of observations of critical opalescence and anomalous viscosity, that solutions in the neighborhood of the critical point might have a colloidal structure. The theory which we have discussed in this paper seems to be in qualitative accord with the idea that the solutions in the critical region have a colloidal or emulsoidal structure, provided they are of the type which have a horizontal portion of the coexistence curve. On the other hand, in the case where the coexistence curve is rounded, if this occurs, the dispersed "ghosts" should have so nearly the properties of the dispersing medium that the solution could hardly be distinguished as a colloidal solution. In this case, however, the classical type of critical opalescence would be expected in the immediate neighborhood of the critical point.

Some solutions do not show the critical opalescence, but Mondain-Monval and Quiquerez (24) showed that this generally happens when the indices of refraction of the liquids involved are nearly alike. They interpreted this as meaning that while the solution was actually colloidal, there were simply no *visible* effects under these circumstances. These authors also investigated the anomalous viscosity of solutions in the critical region (25). They found an anomalously high viscosity in a number of cases in which critical opalescence was not seen. However, they found some ternary mixtures in which there was neither anomalous viscosity nor critical opalescence, and in these cases they supposed that there was molecular dispersion rather than a colloidal solution.

A somewhat more refined analysis can be made if the polarization of light scattered by the liquid is observed. From such measurements Krishnan (16, 17) concluded that molecular clusters of diameter of the order of the wave length were formed near the critical solution temperature. A spongy mass with surfaces having an average radius of curvature of this order of magnitude would probably give a similar effect. An average radius of curvature of the order of the wave length of light seems rather large, however, and it may be that something other than radius of curvature is the controlling factor. There is some evidence, perhaps not too certain, from ultramicroscopic examination (10) and viscosity measurements (27) that the dispersion is finer than the wave length of light.

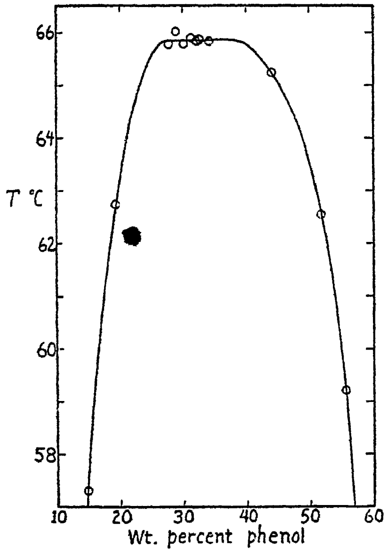


FIG. 2.

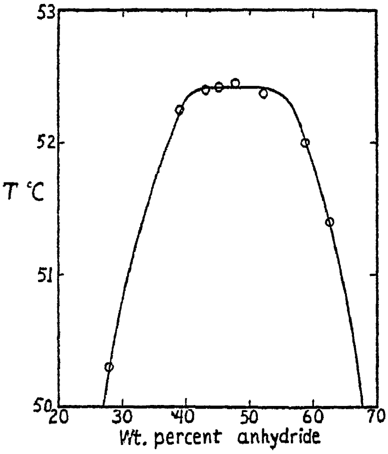


FIG. 3.

FIG. 2. Coexistence curve for phenol and water
FIG. 3. Coexistence curve for acetic anhydride and cyclohexane

The shape of the coexistence curve

If one examines the coexistence curves given in the Landolt–Börnstein tables, or elsewhere, he obtains the impression that many of them are extremely flat at the critical temperature. This has, in fact, been remarked upon by various investigators; for example, Woodburn, Smith, and Tetewsky (39), in the course of an investigation of critical solution temperatures of organic binary mixtures, found that the temperature at which the liquid became homogeneous was independent of concentration over a range of concentrations, and that this aided greatly in the determination of the critical solution temperature. Of course, it is impossible to prove experimentally from data of this sort that there is an absolutely horizontal portion to any coexistence curve; one may always be dealing merely with an extremely flat maximum. However, I think it may be reasonably safely inferred that there are curves with a horizontal portion, and I believe, in fact, that most of the curves do have such a horizontal portion. A couple of illustrative examples are given in figures 2 and 3.

On the other hand, there do seem to be curves which are rounded at the critical temperature. One such example, the best one which I have run across, is shown in figure 4. In most cases, when curves which have the rounded appearance are carefully examined, it appears that the evidence is faulty. Most often the experimental points are not close enough together to enable a decision to be reached. It is, as a matter of fact, just as impossible to exclude a short horizontal section in an apparently rounded curve, such as is shown in figure 4, as it is to exclude the possibility of a very flat maximum in a curve with an apparently completely flat top. As a matter of fact, one might well anticipate much shorter horizontal sections than frequently seem to occur. In the liquid–vapor equilibrium in the case of ethylene, the horizontal section of the coexistence curve extends over a volume range of only 10 per cent. Since the compression of the spongy mass of material at the critical point is mechanically similar in the one-component and the two-component systems, one might suppose that the horizontal section in the two-component system would correspond to a motion of the semipermeable piston considered in Section IV through a volume about 10 per cent of the total volume being compressed. Yet, in many cases, the composition range involved obviously corresponds to a greater volume range. Some rather extreme examples of this are furnished by the systems acetone–carbon disulfide (3) and methyl acetate–carbon disulfide (4), in which cases the flat section ranges from about 50 to about 75 mole per cent carbon disulfide.

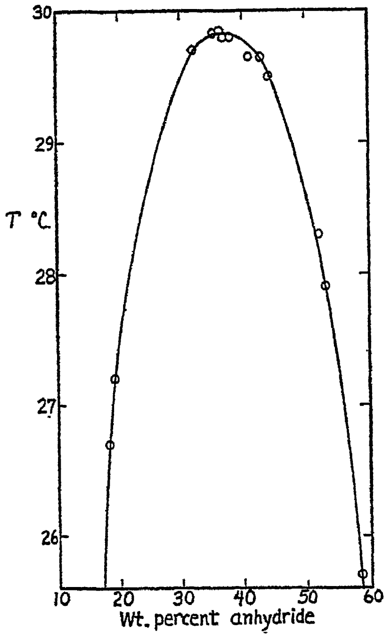


FIG. 4.

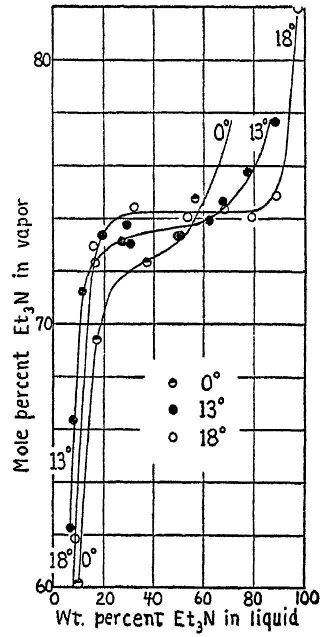


FIG. 5.

FIG. 4. Coexistence curve for acetic anhydride and carbon disulfide

FIG. 5. Vapor–liquid composition curves for triethylamine and water

Another difficulty in being sure that a coexistence curve is actually rounded at the critical point arises from the effect of impurities. It is not unusual for there to be a range of several degrees in the reported values of the critical temperature for a given pair, and the curves reported by different authors may also vary in shape. Small amounts of a surface-active impurity lower the interfacial tension, so that it would become zero before the critical temperature was reached. Emulsification would then result, with increase in the surface, but the amount of material which could be emulsified would depend on the amount of surface-active agent present. Thus the two branches of the coexistence curve might be drawn together, without actually meeting. They would be drawn closer together the lower the original value of the interfacial tension, however; thus a coexistence curve with a rounded top could be produced, even though the curve would be flat if no surface-active agent were present.

The chemical potential in the critical region

There are very few data available which would have a bearing on the question whether the chemical potential can remain constant over a range of concentrations in the region of homogeneous solutions beyond the critical temperature. Probably the best work on this subject is that of Roberts and Mayer (32) on the triethylamine–water system, which has a lower critical temperature at about 18.4°C. These authors measured the mole per cent of amine in the vapor over a range of liquid concentrations. If the chemical potential remains constant over a range of liquid concentrations, the composition of the vapor phase should also be constant over this range. These authors believed that the composition of the vapor did remain constant over a considerable range of liquid concentration at 18°, 16°, and 13°C., while the slope of the curve became definitely finite at 0°C. Of course, a curve which is experimentally of zero slope may actually have a very small slope, and such an experiment can never be absolutely conclusive. However, Roberts and Mayer argued that, since the slopes at 16°C. and 13°C. were very much smaller than would be expected from an interpolation, the slopes must suddenly become larger somewhere between 13°C. and 0°C.; therefore, it is reasonable to suppose that they rise from zero. I do not believe, however, that this argument is really valid. In figure 5 are shown the data of Roberts and Mayer for 18°, 13°, and 0°C. (the data at 16°C. are omitted to avoid confusion), and it will be seen that the curves which are passed through the data fit them reasonably well, do not show any zero slope except at 18°C., just below the critical temperature, and change gradually on going from 18° to 0°C. The slope of the curve at 13°C. is small, but not unexpectedly small if it be compared with the pressure–volume curves of the one-component system (see figure 2 of reference 31). This does not mean that Roberts and Mayer's analysis is necessarily wrong, but I do believe that it indicates that other interpretations are not excluded.

Over a region in which the chemical potentials of the constituents of a binary phase are constant, the total vapor pressure should also be constant. The vapor pressure of aqueous solutions of phenol has been measured by a number of investigators (6, 34, 36, 38). The earlier measurements are not accurate; the accurate measurements of Ferguson (6) are plotted on an exaggerated scale in figure 6. It is seen that there is an almost horizontal section of the curve from about 25 to 45 per cent phenol. We can calculate roughly from the Gibbs–Duhem equation that over this range the partial pressure of water will decrease by about 0.05 mm. and that of phenol will increase by about 0.01 mm. These quantities are zero within the limits of error; nevertheless, we shall assume that they are correct as given, and shall attempt to determine whether they are of a reasonable order of magnitude.

In Section IV we considered a hypothetical experiment involving compression by a semipermeable piston, which was the nearest analogy to compression in a one-component system. Therefore, it appears that the closest analogy to the pressure–volume isotherm for the one-component system would be an osmotic pressure–volume isotherm for the binary system. Let us, therefore, consider how the osmotic pressure Π would change over the range of concentrations from 25 to 45 per cent phenol. Π is given by the equation

$$\Pi = (RT/\bar{v}_a) \ln p_a(0)/p_a(w) \qquad (31)$$

where $\bar{v}_a$ is the partial molal volume of solvent, $p_a(w)$ is its partial vapor pressure at weight per cent w, $p_a(0)$ being the vapor pressure of pure solvent. The partial vapor pressure of phenol over the range $w = 25$ to $w = 45$ per cent may be found from Schreinemakers' measurements (34) (which are not accurate enough to give differences but may be used for rough absolute values of the partial pressures) to be about 6 mm. and that of water is about 284 mm. If then we let $\Delta\Pi$ be the difference in osmotic pressure for $w = 25$ and $w = 45$ per cent, and assume that $\bar{v}_a$ is constant, application of equation 31 gives roughly:

$$\frac{\Delta\Pi}{\Pi} = \frac{\ln (284.05/284)}{\ln (290/284)} = 0.008$$

If equation 31 is applied, considering phenol instead of water as the solvent, we get, using Kahlbaum's (15) value of 10.9 mm. for the vapor pressure of phenol at 75°C.:

$$\frac{\Delta\Pi}{\Pi} = \frac{\ln (6.01/6)}{\ln (10.9/6)} = 0.003$$

As may be seen by comparing with pressure–volume curves for a one-component system just above the critical temperature (see reference 31, Fig. 2) these are not extremely small relative pressure changes, though a change in w from 25 to 45 per cent is equivalent to a rather large volume change; in any event $\Delta\Pi/\Pi$ is subject to considerable error.

To be sure 75°C. is rather far, about 10°C., above the critical temperature for phenol–water. But in this connection we may mention that Antonoff (1) has measured the interfacial tension between the phenol–water phases at 25°C. and finds it to be of the order of 0.5 dyne per centimeter. Since the critical temperature, at which the interfacial tension is presumably zero, is 65°C., it is seen that this gives a rate of change of the interfacial tension with temperature which is very much smaller than that of most surface tensions of pure liquids. Accordingly we may expect the slope of the chemical potential *vs.* concentration curve to change very slowly above the critical temperature.

Effect of addition of a third component

It has already been mentioned that impurities can have an appreciable effect on the coexistence curve of a two-phase system. A great many controlled experiments have been carried out to test the effect of a third component, and it has been found that these effects are often exceedingly large. It is not, perhaps, surprising that Timmermans (37) found that addition of 1 per cent of sodium oleate to the critical mixture of phenol and water lowered the temperature of miscibility by more than 20°C., for sodium oleate is a notoriously surface-active substance. It can lower the already low interfacial tension, so that emulsification will occur, with formation of a single highly disperse phase, at a much lower temperature than would ordinarily be the case. But at first sight it is surprising that a little over 3 per cent of sodium chloride (molal concentration 0.55 per 1000 g. of solution) raises the mixing temperature of this same solution by almost 40°C. (33, 37). This, however, can be explained in the light of the very small temperature coefficient of the interfacial tension noted just above. Because of this small temperature coefficient, a small change in the interfacial tension can cause a large change in the temperature at which it finally becomes zero.

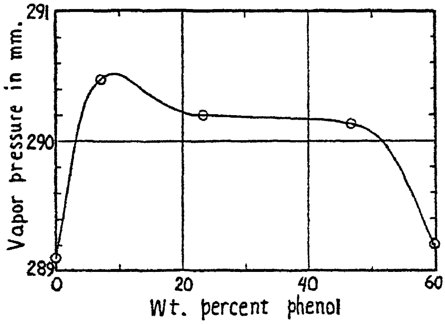


FIG. 6. Vapor pressures of phenol–water solutions at 75° C.

A rough estimate of the order of magnitude may be made by the use of the Gibbs adsorption isotherm,

$$d\gamma_\sigma = -\Gamma_a d\mu_a - \Gamma_b d\mu_b - \Gamma_c d\mu_c \tag{32}$$

where γ_σ is the surface tension, the Γ's are the adsorptions of the various constituents per unit area, and the μ's are chemical potentials. By proper choice of the location of the surface, one of the Γ's can be made equal to zero, and we may reasonably assume that the order of magnitude of the change in γ will be given by retaining only the one term which refers to the "impurity," thus writing:

$$d\gamma_\sigma \sim -\Gamma_c d\mu_c \tag{33}$$

If the impurity is present in small quantities generally $d\mu_c = RT \text{ d} \ln c_c$, where c_c is the concentration; in the case of sodium chloride, however, $d\mu_c = 2RT \text{ d} \ln c_c$, on account of the dissociation, so

$$d\gamma_\sigma \sim -2RT(\Gamma_c/c_c)dc_c \tag{34}$$

In the case of sodium chloride the surface tension is raised, and the adsorption is negative. Let us suppose that there is a region of thickness about 3×10^{-8} cm. near the surface from which sodium chloride is excluded. Then if the concentrations are expressed per cubic centimeter, we have $\Gamma_c/c_c \sim -3 \times 10^{-8}$ cm., $R = 8.3 \times 10^7$ ergs per degree per mole, and $T \sim 350°K$. Hence for a concentration of $c_c = 0.55$ mole per 1000 g. of solvent, $\sim 0.5 \times 10^{-3}$ mole per cubic centimeter, we will find a change in interfacial tension given by integrating equation 34:

$$\Delta\gamma_\sigma \sim 0.9 \text{ dynes cm.}^{-1}$$

In view of the low temperature coefficient of γ_σ, it is clear that this is of the right order of magnitude to raise the mixing temperature by 40°C.

Suggestions for future experiments

It is clear that what is needed now to throw further light on the questions raised in this paper is a thorough and accurate investigation of a few systems with highly purified components, rather than further experiments of a general nature with a large number of systems. Attempts to determine constancy of the chemical potentials by analysis of the composition of the vapor phase seem likely to run into experimental difficulties. Measurements of the total vapor pressure in systems in which the two components have greatly different volatilities and which do not show a vapor pressure maximum seem more promising. Such measurements, however, need to be done with a high degree of accuracy, and differential measurements in which the vapor pressure of a solution of one composition is balanced against that of another composition would be preferable. The surface tension at the liquid–vapor interface is also constant when the chemical potentials are constant. This is another property which could be profitably measured, as was also suggested by Dr. L. D. Roberts in a private communication. Measurements of either total vapor pressures or surface tension should be correlated with observations on the coexistence curve and with measurement of the interfacial tension between the liquid phase as the critical temperature is reached. Similar measurement should be made with a third substance present. It is important that *all* the relevant data should be obtained for a single liquid pair. We are starting a program along these lines in this laboratory, using the cyclohexane–aniline system.

VI. SUMMARY

The coexistence curve in the pressure–volume plane, for the equilibrium between liquid and vapor in a one-component system, is apparently not rounded at the critical point as demanded by the simple van der Waals theory. Instead it shows a horizontal (constant-pressure) portion at the critical temperature. The isotherm at the critical temperature has a horizontal portion, which coincides with the horizontal part of the coexistence curve. According to the theory of Mayer and Harrison, isotherms slightly above the critical temperature should still have a horizontal section. It is the belief of the present author, however, that these isotherms actually have a definite slope, though it is very small near the critical point.

Similar questions can arise in the case of two-component liquid systems. In terms related to the experimentally observed data, these questions are the following: (*1*) Does the coexistence curve for two liquid phases in the temperature–composition diagram show a horizontal (constant-temperature) portion? (*2*) In the temperature region near the critical temperature, but where only one liquid phase exists, can there be a range of concentrations in which the chemical potentials of the components are independent of composition?

In the attempt to answer these questions there are first considered certain thermodynamic properties of the critical solutions. It is shown that if the coexistence curve is rounded, the first and second derivatives of the chemical potentials with respect to concentration must vanish, whereas if the curve has a horizontal part only the first derivative must vanish. Necessary conditions for upper and lower critical temperatures are expressed in terms of the signs of the entropies of transfer of the components from one solution to the other.

The problem is then considered from the point of view of associating molecules. Solute molecules can associate to form clusters. At a certain concentration the very great clusters suddenly become stable, corresponding to the appearance of another phase. This depends, however, upon the interfacial tension between the solution and the large clusters being positive. When the interfacial tension vanishes, the large clusters come in gradually, and at a point where the free volume for large clusters vanishes (concentration infinite in free space) the conditions for a critical point with a horizontal section in the coexistence curve are satisfied. The condition for the vanishing of the interfacial tension is discussed. When the surface tension vanishes the large clusters coalesce to form a spongy mass, a sort of colloid structure. Beyond the critical temperature no meniscus can be seen, but the spongy mass still forms. It has a negative interfacial tension, which resists any attempt to lessen the extent of the surface. It is shown that this leads to the conclusion that there will be no region of constant chemical potential beyond the critical point.

The above paragraph describes the situation if the interfacial tension vanishes because of its temperature coefficient. It is not impossible that there are cases in which the interfacial tension vanishes merely because the two phases finally become identical. This would lead to a rounded coexistence curve.

The experimental evidence bearing on the problem may be briefly summarized as follows: The optical properties and viscosity of solution in the neighborhood of the critical point indicate, as has long been known, that there actually is a sort of colloidal structure in that region, as seems reasonable on the basis of the above discussion. Experimental coexistence curves indicate almost certainly that there are such curves with a horizontal section at the critical temperature. There is some evidence that there are also curves which are rounded in this region. The experimental evidence which bears on the question of the chemical potential consists of one or two measurements on the composition of the vapor phase in equilibrium with the liquid system and a few measurements on total vapor pressures. These data do not answer the question unequivocally, since extremely high accuracy would be required. The effect on the critical temperature of addition of a third component has been considered. It is shown that the experimental results are of a reasonable order of magnitude, by considering the effect of a third substance on the interfacial tension. Finally, a number of suggestions have been made for future experimental work.

VII. REFERENCES

(1) Antonoff, G., Hecht, M., and Chamin, M.: J. Phys. Chem. **45,** 791 (1941).
(2) Band, W.: J. Chem. Phys. **7,** 324, 927 (1939).
(3) Clusius, K., and Ringer, W.: Z. physik. Chem. **187A,** 186 (1940).
(4) Clusius, K., and Ulmke, H.: Z. physik. Chem. **189A,** 331 (1941).
(5) Dacey, J. R., McIntosh, R. L., and Maass, O.: Can. J. Research **B17,** 206 (1939).
(6) Ferguson, J. B.: J. Phys. Chem. **31,** 757 (1927).
(7) Frenkel, J.: Acta Physicochim. U.R S.S. **8,** 261 (1938).
(8) Frenkel, J.: J. Chem. Phys. **7,** 200, 538 (1939).
(9) Friedländer, J.: Z. physik. Chem. **38,** 385 (1901).
(10) Füchtbauer, C.: Z. physik. Chem. **48,** 549 (especially 552) (1904).
(11) Harrison, S. F., and Mayer, J. E.: J. Chem. Phys. **6,** 101 (1938).
(12) Hill, A. E., and Malisoff, W. M.: J. Am. Chem. Soc. **48,** 918 (1926).
(13) Hirschfelder, J., Stevenson, D., and Eyring, H.: J. Chem. Phys. **5,** 896 (and especially 912) (1937).
(14) Jones, D. C., and Betts, H. F.: J. Chem. Soc. **1928,** 1177.
(15) Kahlbaum, G. W. A.: Z. physik. Chem. **26,** 604 (1898).
(16) Krishnan, R. S.: Proc. Indian Acad. Sci. **1A,** 211, 915 (1935); **2A,** 221 (1935); **5A,** 577 (1937); **9A,** 303 (1939).
(17) Krishnan, R. S.: Kolloid-Z. **84,** 2, 8 (1938).
(18) MacDougall, F. H.: *Thermodynamics and Chemistry,* 3rd edition, p. 156. John Wiley and Sons, Inc., New York (1939).
(19) McIntosh, R. L., Dacey, J. R., and Maass, O.: Can. J. Research **B17,** 241 (1939).
(20) McMillan, W. G., Jr., and Mayer, J. E.: J. Chem. Phys. **13,** 276 (1945).
(21) Mason, S. G., Naldrett, S. N., and Maass, O.: Can. J. Research **B18,** 103 (1940).
(22) Mayer, J. E., and Harrison, S. F.: J. Chem. Phys. **6,** 87 (1938).
(23) Michels, A., Blaisse, B., and Michels, C.: Proc. Roy. Soc. (London) **A160,** 358 (1937)
(24) Mondain-Monval, P., and Quiquerez, J.: Bull. soc. chim. **7,** 914 (1940); **11,** 26 (1944).
(25) Mondain-Monval, P., and Quiquerez, J.: Bull. soc. chim. **12,** 380 (1945).
(26) Naldrett, S. N., and Maass, O.: Can. J. Research **B18,** 118 (1940).
(27) Ostwald, Wo: Kolloid-Z. **1,** 331 (and especially 335) (1907).
(28) Ostwald, Wo.: Ann. Physik **36,** 848 (1911).
(29) Ostwald, Wo., and Malss, H.: Kolloid-Z. **63,** 61 (1933).
(30) Raschevsky, N., and Raschevsky, E.: Z. Physik **46,** 300 (1927).
(31) Rice, O. K.: J. Chem. Phys. **15,** 314 (1947).
(32) Roberts, L. D., and Mayer, J. E.: J. Chem. Phys. **9,** 852 (1941).
(33) Schreinemakers, F. A. H.: Z. physik. Chem. **29,** 579 (and especially 597) (1899).
(34) Schreinemakers, F. A. H.: Z. physik. Chem. **35,** 459 (1900).
(35) Schröer, E.: Z. physik. Chem. **A129,** 79 (1927); **A140,** 241, 379 (1929).
(36) Schükarew, A.: Z. physik. Chem. **71,** 90 (1910).
(37) Timmermans, J. C.: Z. physik. Chem. **58,** 129 (1907).
(38) van der Lee, N. J.: Z. physik. Chem. **33,** 622 (1900).
(39) Woodburn, H. M., Smith, K., and Tetewsky, H.: Ind. Eng. Chem. **36,** 588 (1944).

THE VARIATION OF THE PROPERTIES OF ELECTROLYTIC SOLUTIONS WITH DEGREES OF DISSOCIATION[1]

T. F. YOUNG AND L. A. BLATZ

George Herbert Jones Chemical Laboratory, University of Chicago, Chicago, Illinois

Received October 25, 1948

Methods are now available for the determination of degrees of dissociation of electrolytes, both weak and strong, in solutions of finite concentration. Light absorption and Raman spectral methods contribute to a consistent picture of the composition of concentrated nitric acid solutions. The Debye–Hückel theory, conductance data, kinetic measurements, and the intensities of Raman lines give a reasonably consistent picture of the molecular constitution of aqueous sulfuric acid in solutions of all concentrations.

The data now are sufficient to permit calculation of reasonably satisfactory theoretical curves showing the heats of dilution, apparent molal volumes, and surface tensions of sulfuric acid solutions. Data are insufficient for an entirely satisfactory comparison of computed and experimental values of heat capacities, but the calculations agree as well as can be expected with the uncertain experimental data.

When the Debye–Hückel theory (9) first appeared there were very few data which were sufficiently accurate for testing it. One of the most interesting and significant tests would have been a comparison of the predictions of the theory with precise experimental values of the degrees of dissociation of weak and strong electrolytes. Although the major objective of the vast amount of work that has been done on electrolytic dissociation was the determination of equilibria in solutions of finite concentration, no adequate determinations existed in 1923.

The answers to a number of other fundamental problems of physical chemistry have had to await more knowledge of electrolytic dissociation. Enough information has now accumulated to permit us to make significant progress with one group of problems: namely, the explanation in terms of constitution of variations of familiar properties of electrolytes with concentration and temperature. Some of the most interesting properties—heat of dilution, density, surface tension, and heat capacity—are dealt with in this review.

I. DETERMINATION OF DEGREES OF DISSOCIATION OF ELECTROLYTES

A. *Conductance methods*

In 1926 two attempts were made to obtain better estimates of degrees of dissociation from conductance data. MacInnes (30) described a method for the investigation of electrolytes of the 1-1 valence type which he has continued to use and develop (31, 33). Prior to 1926 the degree of dissociation had been calculated from the ratio of Λ, the equivalent conductance, to Λ_0, the limiting equivalent conductance at infinite dilution. MacInnes employed instead:

---

[1] Presented at the Symposium on Thermodynamics and Molecular Structure of Solutions, which was held under the auspices of the Division of Physical and Inorganic Chemistry at the 114th meeting of the American Chemical Society, Portland, Oregon, September 13 and 14, 1948.

$$\alpha = \Lambda/\Lambda_e \tag{1}$$

where Λ_e is the equivalent conductance of only the dissociated part of the electrolyte. Λ_e, in principle, could be determined from Λ_0 by the application of corrections for the influence of interionic forces and other similar effects. In the procedure used by MacInnes, the estimate of Λ_e was based upon the assumption that the sum of the mobilities of the ions of a uni-univalent electrolyte is determined by the total ionic strength. He also used the Debye–Hückel limiting law for estimation of activity coefficients to be used in his calculations. The apparent equilibrium constants obtained with this procedure are essentially constant for acetic acid concentrations ranging as high as 0.0002 mole per liter (32). That concentration is still small, but the improvements in methods and results were definite and important.

In the same year Sherrill and Noyes (55) attacked a somewhat more difficult problem: they dared to work with ions of mixed valence types and at concentrations as high as 0.05 molal. The procedure they developed is, as a result, considerably less general than that of MacInnes. They assumed that the mobility of an ion is the same in two solutions of the same ionic strength, provided the ions in each solution are similar (in structure, valence, etc.). There was little they could do at the time to demonstrate the validity of their method, but apparently they presented their results with considerable confidence. Some of their conclusions concerning sulfuric acid solutions will be compared below with the results of methods developed more recently.

B. *The first optical method: light absorption*

von Halban and Brüll (13) have recently compared and appraised various methods which have been employed repeatedly in investigations of the dissociation of weak and moderately strong electrolytes. They reiterated earlier conclusions of von Halban and his coworkers that the method of light absorption stands out above the other well-known procedures because it is not highly dependent upon corrections for the influences of solute ions and neutral molecules. The conductance method involved assumptions concerning the variations of ionic mobilities which are large even when solutions are relatively dilute. Freezing-point, vapor pressure, osmotic pressure, and electromotive force methods are affected by assumptions concerning relations of activities and concentration. Some, at least, of the electromotive force methods are marred also by errors due to liquid-junction phenomena for which adequate corrections have not been devised. The molecular extinction coefficient of an absorbing ion, however, is sensibly constant for all ionic strengths less than 0.01. In an earlier paper von Halban and Eisenbrand (14) had observed large changes in the molecular extinction coefficient when the ionic concentrations became very large, but at low concentrations the variations are negligible. With refined optical methods, such as those employed by von Halban and Kortüm (15, 26), a precision as good as

0.01 per cent can be obtained in the measurement of extinction coefficients, and as good as 0.1 per cent in the determination of dissociation constants.

A severe and beautiful test of the Debye–Hückel theory made in 1934 by von Halban and Kortüm (15) is illustrated in figure 1. The ordinates represent the logarithms of the apparent dissociation constant of 2,4-dinitrophenol. The apparent constant (the function calculated directly from concentrations without the aid of activity coefficients) should vary in dilute solutions, as indicated by the straight line whose slope was calculated by means of the Debye–Hückel limiting law. Some of the points at very low concentrations represent solutions containing only the solute, dinitrophenol. Other points show the effects of added neutral salts. The Debye–Hückel limiting law is obeyed almost perfectly for all ionic strengths less than 0.003. It fails, as expected, at higher concentrations, but the principle of ionic strength remains valid to an ionic strength of 0.01. The authors extended their observations to ionic strengths greater than 2.00. In solutions of those high concentrations they observed large specific ion effects. It is not yet possible to determine how much of such an effect is due to influences which alter the degree of dissociation, and how much should be ascribed to variations in the molecular extinction coefficient itself. The usefulness of the method for studying concentrated solutions is therefore limited.

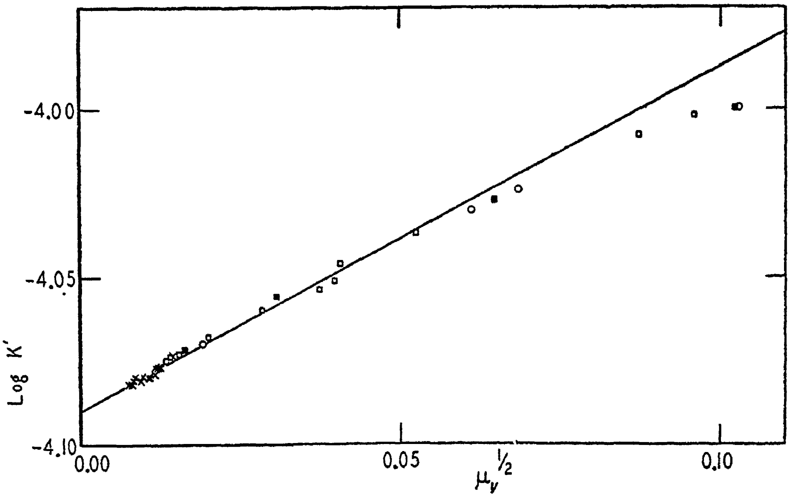


FIG. 1. Logarithm of apparent dissociation constant of 2,4-dinitrophenol at 25°C × indicates acid as sole solute; ○, acid in presence of potassium chloride; □, acid in presence of potassium perchlorate; and ■, acid in presence of sodium chloride. Straight line represents Debye–Hückel limiting slope.

For very dilute solutions, the most important disadvantage of the light absorption method is that it is restricted to equilibria which involve at least one ion or molecule that absorbs light. Though colorless acids or bases can sometimes be studied with the aid of indicators, the procedure is more complex and less accurate.

C. The second optical method: Raman spectra

Another optical method which makes use of Raman radiation has been developing during the last fifteen or twenty years. The method depends upon measurements of the relative intensities (usually determined photographically) of a Raman line produced by exciting radiation supplied by a source of constant strength. Application to investigations of chemical equilibria has usually depended upon the assumption that the intensity of a selected Raman line is proportional to the concentration of the ion or other molecule giving rise to that line. The accuracy of the simple proportionality has been investigated for only a very few substances and inadequately for those. The special attention directed toward nitric acid, however, has resulted in relatively extensive investigations of the line of $\Delta\tilde{\nu} = 1048$ cm.$^{-1}$ arising in solutions of various nitrates. Sodium nitrate solutions have been studied by I. R. Rao (39), Simons (56), and N. R. Rao (41). Simons also studied ammonium nitrate solutions. Relative intensities for these nitrates, taken from the tables of Simons and of N. R. Rao, are shown in figure 2. Simons' measurements of intensities of the 980 cm.$^{-1}$ line from ammonium sulfate solutions are also plotted. The scales used in figure 2 for the intensities are entirely arbitrary. They were chosen, merely to avoid superposition of the lines, so that the slopes would be 0.75, 1.00, and 1.25. Obviously the ratio of intensity to concentration is constant, within the limits of precision of the experimental methods, throughout the ranges of concentration investigated.

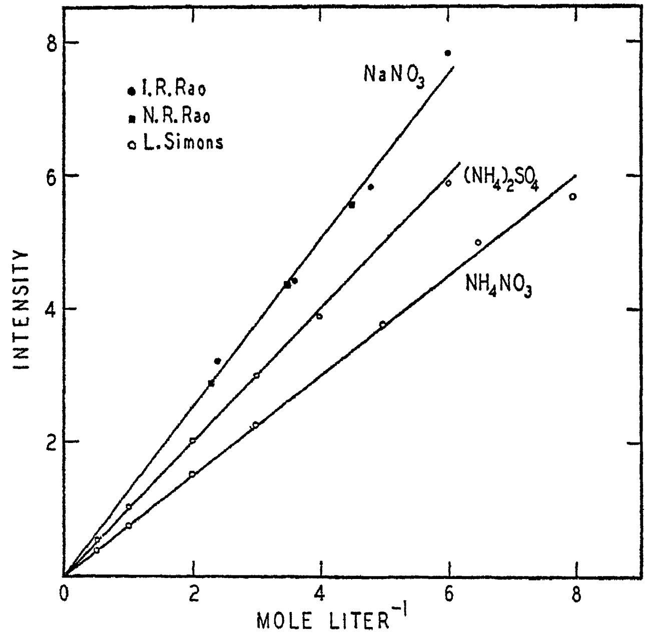


FIG. 2. Relative intensites of Raman lines from solutions of sodium nitrate, ammonium nitrate, and ammonium sulfate. Vertical scales are purely arbitrary.

(1) Nitric acid

The intensity of the Raman line of 1048 cm.$^{-1}$ arising in a concentrated nitric acid solution is not so great as its intensity when produced in a sodium nitrate solution of the same concentration. Moreover, when the concentration of the acid is increased beyond 39 weight per cent the intensity of the 1048 cm.$^{-1}$ line actually decreases. Simultaneously the intensities of other lines (notably the line of 1292 cm.$^{-1}$) characteristic of pure nitric acid become greater.[2]

Redlich and Bigeleisen (46) determined experimentally the concentration of the sodium nitrate solution giving rise to lines of the same intensity as each of several nitric acid solutions. Equating the ratio of the concentrations to the degree of dissociation of the acid, they obtained the results[3] in table 1. From these values they estimated the thermodynamic dissociation constant ($K = 21 \pm 4$). Because the present knowledge of activity coefficients is very limited, the uncertainty in the constant is large. If, however, the assumptions that sodium nitrate is completely dissociated and that the nitrate-ion concentration is the same in solutions giving rise to radiations of equal intensity are accurate, the uncertainties in the degrees of dissociation in table 1 are no greater than those in the experimental work itself.

TABLE 1

Degree of dissociation of nitric acid

C...........................	4.51	6.60	8.90	10.30	11.89	14.23
α...........................	0.823	0.673	0.49	0.39	0.32	0.14

An interesting application of the data in table 1 may be made to the interpretation of the dependence on concentration of the molecular extinction coefficients of nitric acid. von Halban and Eisenbrand (14) have determined curves for light of several wave lengths, showing the variation of the molecular extinction coefficients with concentration. The curves are complex, containing prominent points of inflection, whereas corresponding curves for potassium nitrate (14, 54) are comparatively simple. The points of inflection may readily be ascribed to the rapid rise in the ratio of the concentration of the undissociated nitric acid to that of the nitrate ion as the total concentration increases. From the known degrees of dissociation and the light absorption data for solutions of potassium nitrate and of nitric acid, it is possible to calculate approximately the molecular extinction coefficients of the undissociated nitric acid. The values of the coefficient so calculated exhibit a maximum at about 2800 Å. and then decrease toward zero as the wave length increases beyond 3100 Å.

---

[2] The assignment of the various observed frequencies (about 10) in HNO_3 and DNO_3 to the expected modes of vibration of the nitric acid molecule has been successfully carried out (*cf.* Redlich and Nielsen (48)).

[3] Values of α for lower concentrations and references to the extensive earlier literature have been given by Redlich (44). His article is an excellent review of the general problem of the dissociation of strong electrolytes.

Dalmon (7) has measured the absorption of light of 3180 Å. in solutions of nitric acid varying in concentration from 0.8 to 24.6 moles per liter (pure acid). Since undissociated nitric acid does not absorb a significant fraction of light of that wave length, his measurements yield directly approximate relative values of the nitrate-ion concentration and therefore of the degree of dissociation of nitric acid. The results are not strictly comparable to those of Redlich and Bigeleisen, because of the known variations with concentration of the molecular extinction coefficients (14) and because Dalmon's work was done at about 1°C. rather than 26°C. Nevertheless, the fact that Dalmon's data lead to degrees of dissociation almost proportional to those of Redlich and Bigeleisen is an effective argument for the essential correctness of the data shown in table 1.

(2) Other electrolytes

Solutions of only a few other electrolytes have been extensively investigated (19). Iodic acid (42), perchloric acid (47), and trichloroacetic acid (43) have all been shown to be incompletely dissociated strong electrolytes. Mixtures of nitric acid and sulfuric acid and of sulfur trioxide and sulfuric acid have also been studied extensively. For our present purposes, the most interesting system is that of sulfuric acid and water, because its properties exhibit marked variations with temperature and concentration which are more complex and interesting than those of most other electrolytes. These phenomena are due to the favorable magnitude of the dissociation constant of the bisulfate ion.

(3) Sulfuric acid solutions

Figure 3 is a simplified representation of various classes of lines reported for solutions of sulfur trioxide in water. The line of 1271 cm.$^{-1}$ is typical of a group of lines arising in pure sulfur trioxide. That line and about fifteen others are probably due to a non-planar molecule with the formula S_3O_9. Other lines seem to be due to the monomer and, at low temperature, to other polymers (10, 11, 18). As water is added to sulfur trioxide, the intensity of the 1271 cm.$^{-1}$ line drops rapidly while other lines appear and increase in intensity (6, 36). Note the very strong 910 cm.$^{-1}$ line whose intensity rises to a maximum in solutions having the composition $SO_3 \cdot H_2O$. As more water is added the intensity of the 910 cm.$^{-1}$ line drops and a line of 1043 cm.$^{-1}$ appears, increases to a maximum, and finally decreases toward zero (40). Another important line, of 980 cm.$^{-1}$, reaches a maximum in a solution of about 15 mole per cent sulfur trioxide. Another, of 960 cm.$^{-1}$, reaches a maximum at a mole fraction of sulfur trioxide somewhere between 60 and 70 mole per cent (36), but the existing data are too meager to indicate accurately the height or position of the maximum. The line has, however, been studied carefully in solutions more dilute than 50 mole per cent sulfur trioxide (2) and more concentrated than 80 mole per cent (36). In both of these regions the intensity has been shown to be low—nearly zero.

A complex literature has been summarized and greatly simplified in figure 3. Only five typical lines of the many reported are represented. Furthermore, each line is identified by a single wave number. Actually the lines have breadth, i.e., extend over a short range of wave lengths. The shapes of the lines, notably the positions of the maxima, vary with composition (1, 2, 35, 59). Most observers have not worked carefully enough to be forced to take the shape and exact position of the lines into account. The intensities represented in the right half of the diagram were estimated only roughly on a whole number scale from 0 to 10 (36). Fortunately for our purposes, the data shown in the left half of the figure are somewhat more accurate. They were obtained by N. R. Rao (40), using a relatively precise photographic procedure. The intensities he reported are actually the integrated intensities obtained from measurements over the full widths of the lines. He assumed that it is the integrated intensity which is proportional to concentration. His results seem to be accurate to about 10 per cent. Unfortunately, Rao did not report the temperatures of his solutions and presumably did not measure them. He reported only that they were cooled (by water) during the exposures.

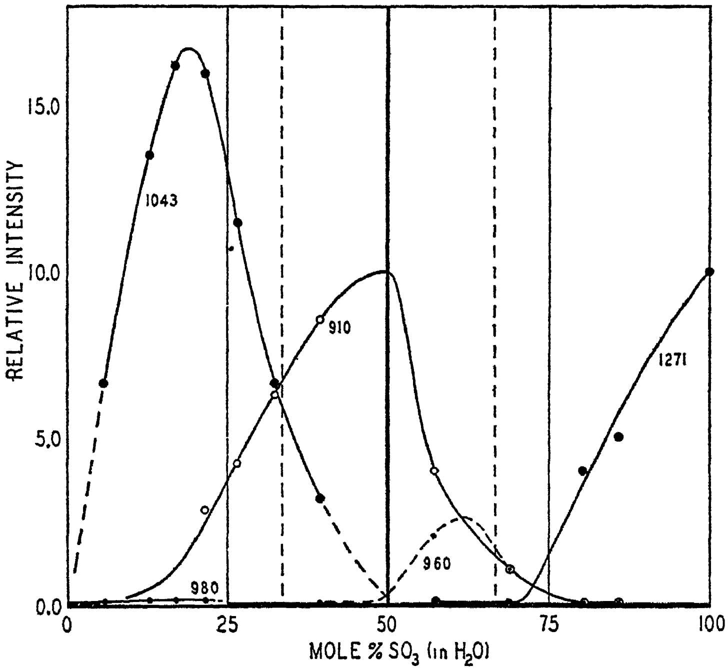


Fig. 3. Intensities of typical Raman lines from solutions of sulfur trioxide and water

Though additional information can yet be derived from more extensive and more precise work, the interpretation of figure 3 seems to be clear. Addition of SO_3 to H_2SO_4 or of H_2O to SO_3 produces pyrosulfuric acid, $H_2S_2O_7$, and the ions $HS_2O_7^-$ and $S_2O_7^{--}$. Presumably the 960 cm.$^{-1}$ line therefore is due to pyrosulfuric acid or one of the products of its dissociation (6). It is interesting to study the decrease in intensity of the 910 cm.$^{-1}$ line as sulfur trioxide is added to sulfuric acid. If we assume that pure sulfuric acid solution contains the one molecular species H_2SO_4, and that 1 mole of it is removed by each gram formula weight of sulfur trioxide added, we can compute the reduction in the concentration of H_2SO_4. These assumptions and the density data of Knietsch (24) lead to the conclusion that the concentration of H_2SO_4 in a solution of 57 mole per cent sulfur trioxide is about 55 per cent of that in pure sulfuric acid. This result is in satisfactory agreement with the 40 per cent indicated on the semiquantitative curve in figure 3. More careful work will probably add considerably to knowledge of the equilibrium in mixtures of sulfur trioxide and water.

The region between 0 and 50 mole per cent sulfur trioxide represents solutions of H_2SO_4 in water. Because he was interested in the relative concentrations of the ions, N. R. Rao measured also the intensities of lines from ammonium sulfate and potassium bisulfate solutions. The strength of the 980 $cm.^{-1}$ line produced in an ammonium sulfate solution and that of the 1043 $cm.^{-1}$ line from a potassium bisulfate solution support the conclusions drawn by Rao and previous workers (2, 35, 59) that the 910, 1043, and 980 $cm.^{-1}$ lines are produced by molecular H_2SO_4, HSO_4^-, and SO_4^{--}, respectively. The usual assumptions (that ammonium sulfate is completely dissociated, and that the ratio of the intensities of the 980 $cm.^{-1}$ line from a sulfuric acid solution and from the ammonium sulfate solution is equal to the ratio of the concentrations of the sulfate ion in the two solutions) lead directly to calculations of the sulfate-ion concentrations in the sulfuric acid and potassium bisulfate solutions[4] (see table 2, column 3).

Subtraction of 0.186 from 3.0 leaves the concentration of bisulfate ion in the 3 molar potassium bisulfate solution. Simple proportion and the measured intensities of the 1043 $cm.^{-1}$ line then yield the concentration of the bisulfate ion in the various sulfuric acid solutions (see column 5). The sum of the SO_4^{--} and HSO_4^- concentrations subtracted from the stoichiometric concentration leaves the concentration of the undissociated H_2SO_4 (column 7). A severe test can be applied to the whole calculation. The 910 $cm.^{-1}$ line should be proportional to the sulfuric acid concentration in column 7. Figure 4 shows that the relation between intensity and the calculated concentration is indeed linear within the limits of error of the photographic technique.

The apparent equilibrium constants K_1' and K_2' were computed for the first and second dissociations directly from the concentrations in table 2. The values of K_2' are compared in figure 5 with the thermodynamic constant determined by Young, Klotz, and Singleterry (22, 57, 61) by means of a light absorption method used with an indicator. The short straight line shows the limiting tangent calculated from the Debye–Hückel theory; the short curved line represents the Debye–Hückel approximation calculated according to the method of Kielland (21) with ion parameters of 9, 4, and 4 Å. for the hydrogen ion, sulfate ion, and bisulfate ion, respectively.

---

[4] Rao's assumptions were equivalent to these, but he did not express his results in terms of actual concentrations. The calculations he made were not explained and contained a sufficient number of arithmetic mistakes to make them difficult to follow. They are also awkward for our purposes. All of the calculations described here were based upon the original integrated intensities given in his article. Values similar to those in column 7 were calculated also by De Béthune and Kimball (8).

The upper curve of figure 5 shows values of K_1'. Since the data are insufficient for extrapolation, the curve has been extended arbitrarily to a value of K_1 estimated roughly from the theory of Kossiakoff and Harker (27). The vertical length of the lines represents the uncertainty introduced by a 10 per cent uncertainty in one of the intensity measurements.

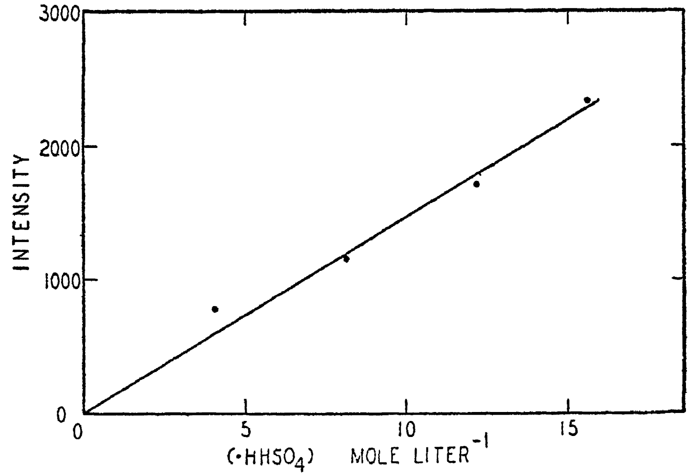


FIG. 4. Intensity of Raman line of 910 cm.$^{-1}$ *versus* concentration of undissociated sulfuric acid calculated by difference.

TABLE 2

Relative intensities of Raman lines and concentrations calculated from them

(1) (H_2SO_4)	(2) 980 cm.$^{-1}$	(3) (SO_4^{--})	(4) 1043 cm.$^{-1}$	(5) (HSO_4^-)	(6) 910 cm.$^{-1}$	(7) ($\cdot HHSO_4$)
16.9			860	1.28	2330	15.6
14.9			1800	2.68	1700	12.2
12.8			3120	4.65	1150	8.2
10.8	11.8	0.25_2	4350	6.48	765	4.1
8.7	15.2	0.32_6	4412	6.57		1.8
6.8	13.0	0.27_8	3678	5.48		1.0
3.0	6.5	0.13_9	1800	2.68		0.2
$(NH_4)_2SO_4$						
3.0	140.5	(3.00)				
$KHSO_4$						
3.0	8.7	0.18_6	1890	2.81		

The fractions of the solute present as sulfate ions, as bisulfate ions, and as undissociated sulfuric acid are represented in figure 6 by the vertical distances between adjacent lines. The sum of the distances represents unity at all concentrations. In very dilute solutions nearly all of the sulfuric acid is present as sulfate and hydrogen ions; at higher concentrations most of it is bisulfate and hydrogen ions; and in very concentrated solutions most of it is undissociated sulfuric acid.

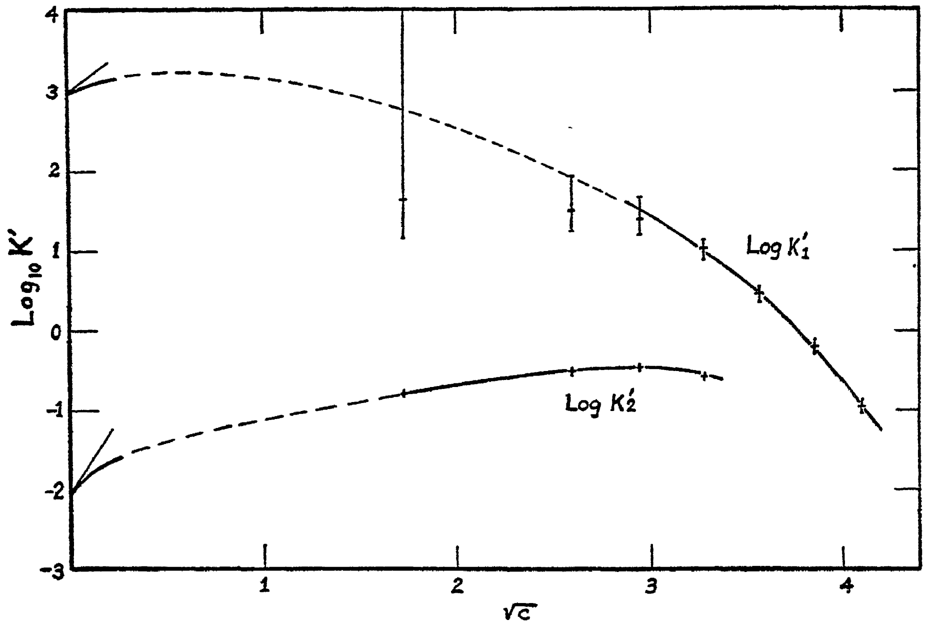


FIG. 5. Apparent dissociation constants K_1' and K_2' of sulfuric acid in water. Vertical lines indicate changes that would be produced in K' by alterations of 10 per cent in measured intensities.

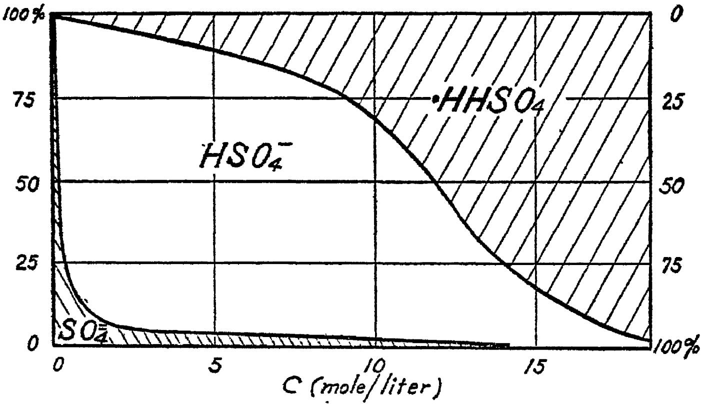


FIG. 6. Fractions of sulfuric acid present as undissociated acid, as bisulfate ion, and as sulfate ion, represented by vertical distances between adjacent lines.

Absolute concentrations may also be computed. In table 3 are shown approximate maximum values of the concentration of the sulfate ion and bisulfate ion, and of the volume and weight ionic strengths. In the bottom line are approximate values of the concentrations of sulfuric acid in which the maxima are attained. The hydrogen-ion concentration reaches a maximum slightly higher, of course, than the maximum (HSO_4^-).

Rao did not investigate solutions more concentrated than 16.9 moles per liter. It is impossible, therefore, to estimate directly from his intensity data the concentration of bisulfate ion in the pure acid. Extrapolation indicates, however, that it is very small. This result is in agreement with the conclusion of Hammett and Deyrup (16) that the bisulfate-ion concentration is about 0.05 mole per liter at 25°C.

For calculations to be described below, it is necessary to calculate degrees of dissociation of the bisulfate ion for a wide range of concentrations and for temperatures varying from 0° to 50°C. Both the dissociation constant and the standard free energy of dissociation vaiy too rapidly with temperature to be treated as constant. A better approximation is to treat activity coefficients as temperature invariants. To facilitate calculations of equilibria in dilute solutions of sulfuric acid, Bray and Liebhafsky (4) prepared a table of equilibrium quotients which are nearly equivalent to γ_B, the product of the activity coefficients of the hydrogen ion and sulfate ion divided by the activity coefficient of the bisulfate ion.

TABLE 3

Maximum concentrations and maximum ionic strengths

	(SO_4^{--})	(HSO_4^-)	μ_V	μ_W
Maximum value..	0.35	6.7	8.0	14.3
Concentration of solution at which maximum occurs........... ...	8–10	10	10	11.5

$$\gamma_B = \frac{(\gamma_{H^+})(\gamma_{SO_4^{--}})}{(\gamma_{HSO_4^-})} \tag{2}$$

They based their table upon the degrees of dissociation determined by Sherrill and Noyes (55), supplementing them with kinetic data of their own. Values of γ_B calculated from their paper are shown in figure 7. Shown also are values of γ_B computed from the Raman data of table 2, and two curves calculated from the Debye–Hückel limiting law (labelled D. H. L.) and from the Debye–Hückel approximation (labelled D. H. A). The D. H. A. curve, calculated by the method of Kielland, was based upon ion parameters of 9, 4, and 4 Å. for the hydrogen ion, sulfate ion, and bisulfate ion, respectively.

On the right portion of the plot two series of points are represented: one group calculated from ion concentrations expressed in molarity (moles per liter of solution) and one group expressed in molality (moles per kilogram of solvent). The former are plotted against the square root of the volume ionic strength, μ_V, and the latter *versus* the square root of weight ionic strength, μ_W. By accident, most of the points appear to lie on the same curve. A point lying above the curve at $\sqrt{\mu_V} = 2.6_9$ represents a considerably more concentrated solution than that at $\sqrt{\mu_V} = 2.7_4$, but appears at the left of the latter because of the maximum attained by the volume ionic strength.

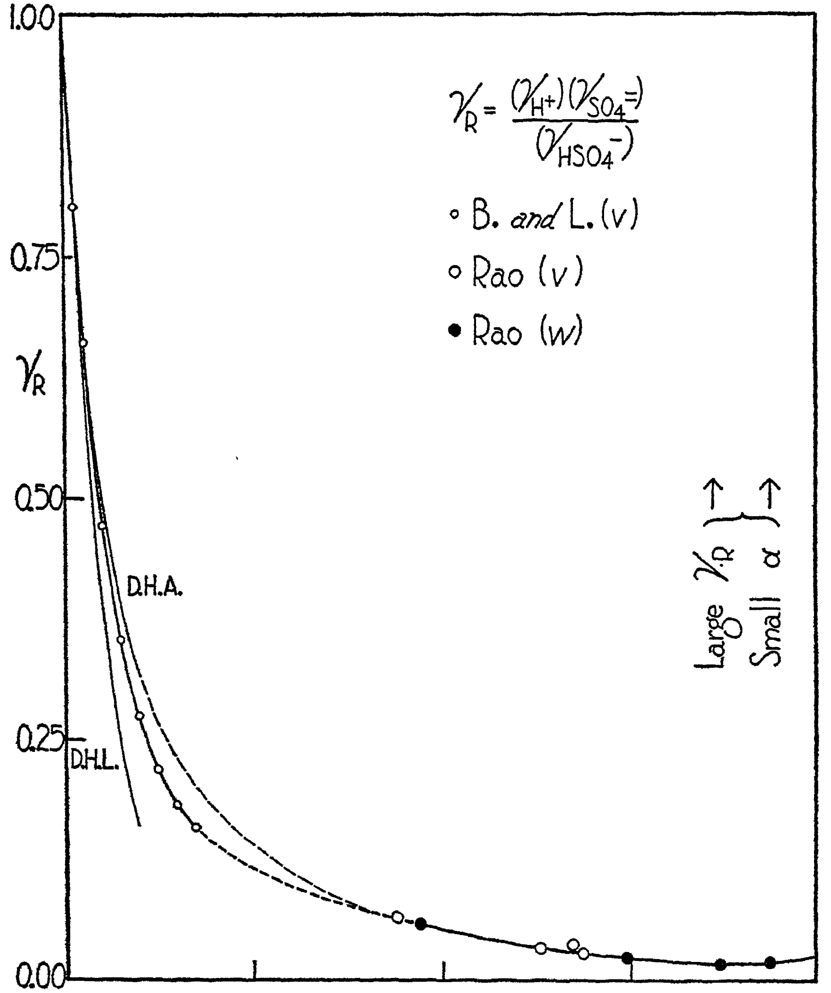


FIG. 7. The activity function γ_R from Raman data, from tables of Bray and Liebhafsky, from the Debye–Hückel limiting law (D. H. L.), and from the Debye–Hückel approximation (D. H. A.).

The agreement between the two sections of the curve is fairly satisfactory and is all that could be expected at the present time. They have been joined by a broken line. Except at very low concentrations, the uncertainties are 10 per cent or more. The Raman data, for example, were treated as 25°C. measurements, although the temperature was not reported. In addition, experimental errors may be as large as 10 per cent. The conductance theory used by Sherrill and Noyes and the methods of Bray and Liebhafsky also may introduce errors of considerable magnitude.

Because of those uncertainties, the experimental curve is probably no more accurate than the theoretical one (D. H. A.) at ionic strengths of 0.2 or less. The latter curve has accordingly been connected by a second broken line with the section representing the Raman data. Within a few years, improvements in the Raman data should make it possible to determine which of the two curves is more accurate. For the present there is little basis for a choice between them for the calculations to follow.

APPLICATION OF ACTIVITY COEFFICIENT CURVE TO ANALYSES OF PROPERTIES OF SOLUTIONS OF SULFURIC ACID

Throughout the following discussions aqueous solutions of sulfuric acid are regarded as mixtures of solvent and three solutes, two of which are completely dissociated. To distinguish the three solutes, the following special notations are used: The formula $H \cdot H \cdot SO_4$ represents that portion of the sulfuric acid which is dissociated into hydrogen ions and sulfate ions; the symbol $H \cdot HSO_4$ denotes the other ionized solute; and $\cdot HHSO_4$ denotes undissociated acid. The concentration of $H \cdot H \cdot SO_4$ is simply the sulfate-ion concentration, and the concentration of $H \cdot HSO_4$ is the concentration of bisulfate ion. The customary formula H_2SO_4 denotes, as usual, the mixture of ionic and molecular species present at equilibrium in a sulfuric acid solution. Obviously:

$$(H_2SO_4) = (H \cdot H \cdot SO_4) + (H \cdot HSO_4) + (\cdot HHSO_4) \tag{3}$$

where the parentheses as usual indicate the respective concentrations.

The symbol φ stands for any apparent molal quantity (29). φ_H, φ_V, and φ_C denote the apparent molal heat content, apparent molal volume, and apparent molal heat capacity, respectively. φ_C at constant pressure only is needed; hence the subscript p is not used. φ^0 represents φ of a substance in its standard state, and ΔH^0 represents increases in heat content (enthalpy) for reactions involving substances all of which are in their standard states. Thus, φ_V^0 $(H \cdot HSO_4)$ represents the apparent molal volume of $H \cdot HSO_4$ in its standard state. When the difference between φ^0 and φ is written, the appropriate formula is placed below the pair of symbols, instead of in parentheses after each φ.

A. Heat of dilution

Heat is evolved slowly and regularly as a mole of sulfuric acid is diluted from 1 molal toward 0.05 molal (5). As it is diluted beyond 0.05 molal, much larger amounts of heat are evolved. That portion of the dilution curve has been investigated at 25°C. by Lange, Monheim, and Robinson (28; *cf.* 17, p. 439), with whose data our theoretical calculation may be compared.

Since we are concerned with dilute solutions only, we can neglect the presence of $\cdot HHSO_4$. A mole of H_2SO_4 is therefore regarded simply as a mixture of α moles of $H \cdot H \cdot SO_4$ and $1 - \alpha$ moles of $H \cdot HSO_4$. The apparent molal heat content of H_2SO_4 is then $\alpha\varphi_H$ of $H \cdot H \cdot SO_4$ plus $(1 - \alpha)\varphi_H$ of $H \cdot HSO_4$ plus ΔH_M, the heat absorbed during the mixing of hypothetical solutions of $H\cdot H\cdot SO_4$ and $H\cdot HSO_4$.

$$\varphi_H(H_2SO_4) = \alpha\varphi_H(H\cdot H\cdot SO_4) + (1 - \alpha)\varphi_H(H\cdot HSO_4) + \Delta H_M \qquad (4)$$

At the concentrations of importance here we shall be able to neglect ΔH_M when the ionic strength of each of the two hypothetical solutions is the same as that of the mixture existing in the real sulfuric acid solution.[5]

The heat of dilution of 1 mole of sulfuric acid from molality m to infinite dilution is the difference between $\overset{0}{\varphi}_H$, the standard apparent molal heat content, and φ_H of sulfuric acid at molality m. Thus:

$$\Delta H \text{ (dilution to } m = 0) = \overset{0}{\varphi}_H - \varphi_H \qquad (5)$$

Now $\overset{0}{\varphi}_H$ of H_2SO_4 and $\overset{0}{\varphi}_H$ of $H\cdot H\cdot SO_4$ are equal, since the conventional standard for each is the infinitely dilute solution.

$$\overset{0}{\varphi}_H(H_2SO_4) = \overset{0}{\varphi}_H(H\cdot H\cdot SO_4) \qquad (6)$$

The dilution of sulfuric acid may be thought of as the sum of four hypothetical processes. Consider, for example, the dilution of 1 mole of H_2SO_4 in 0.01 molal solution to infinite dilution. The first step is the separation of a solution containing 1 mole of H_2SO_4 and 10^5 g. of water into two portions containing α mole of $H\cdot H\cdot SO_4$, and $1 - \alpha$ mole of $H\cdot HSO_4$, respectively. The amount of solvent in each is sufficient to make each ionic strength equal to that of the original solution, which is $\mu_W = m(1 + 2\alpha) = 0.01(1 + 2\alpha)$.

$$\begin{array}{c} \text{Solution of } H_2SO_4 \\ \text{and } 10^5 \text{ g. } H_2O \end{array} \rightarrow \begin{array}{c} \text{Solution of} \\ \alpha H\cdot H\cdot SO_4 \text{ and} \\ \dfrac{3\alpha}{1 + 2\alpha} 10^5 \text{ g. } H_2O \end{array} + \begin{array}{c} \text{Solution of} \\ (1 - \alpha)H\cdot HSO_4 \text{ and} \\ \dfrac{(1 - \alpha)}{1 + 2\alpha} 10^5 \text{ g. } H_2O \end{array} \qquad (7)$$

ΔH of this separation process is the negative of ΔH_M at constant ionic strength.

Now to each hypothetical solution is added an infinite amount of water:

$$\begin{array}{c} \text{Solution of} \\ \alpha H\cdot H\cdot SO_4 \text{ and} \\ \dfrac{3\alpha}{1 + 2\alpha} 10^5 \text{ g. } H_2O \end{array} + \infty H_2O \rightarrow \begin{array}{c} \text{Solution of} \\ \alpha H\cdot H\cdot SO_4 \text{ and} \\ \infty H_2O \end{array} \qquad (8)$$

---

[5] A series of investigations now in progress shows that the changes in total volume and in heat content accompanying the mixing of aqueous solutions of electrolytes of the 1-1 valence type are very small when each solution has the same ionic strength as the ternary mixture produced. For solutions as dilute as those involved here, the same is true of the mixing of electrolytes of the 1-1 valence type, e g., sodium chloride, and of the 1-2 valence type, e.g., potassium sulfate (58). There is some indication, however, in current investigations of the free energy of mixing that changes in volume, heat content, and excess free energy may be even smaller if all of the solutions are of the same normality (49). For the concentrations involved in this calculation, either formulation of the principle leads to essentially the same results.

$$\begin{array}{l} \text{Solution of} \\ (1-\alpha)\mathrm{H\cdot HSO_4} \text{ and} \\ \dfrac{(1-\alpha)}{1+2\alpha}\, 10^5 \text{ g. } \mathrm{H_2O} \end{array} + \infty \mathrm{H_2O} \rightarrow \begin{array}{l} \text{Solution of} \\ (1-\alpha)\mathrm{H\cdot HSO_4} \text{ and} \\ \infty \mathrm{H_2O} \end{array} \qquad (9)$$

Finally, the $(1 - \alpha)$ mole of infinitely dilute $\mathrm{H \cdot HSO_4}$ is allowed to dissociate:

$$(1 - \alpha)\mathrm{H \cdot HSO_4} \rightarrow (1 - \alpha)\mathrm{H \cdot H \cdot SO_4} \qquad (10)$$

ΔH of this process is $(1 - \alpha)[\varphi_H^0$ of $\mathrm{H \cdot H \cdot SO_4} - \varphi_H^0$ of $\mathrm{H \cdot HSO_4}]$ which is $(1 - \alpha)$ times the standard heat of dissociation, $\Delta H_{\mathrm{Dis}}^0$, of the bisulfate ion. The heat of dilution of sulfuric acid from any m to $m = 0$ is therefore:

$$\underset{\mathrm{H_2SO_4}}{\varphi_H^0 - \varphi_H} = -\Delta H_M + \alpha\underset{\mathrm{H \cdot H \cdot SO_4}}{(\varphi_H^0 - \varphi_H)} + (1 - \alpha)\underset{\mathrm{H \cdot HSO_4}}{(\varphi_H^0 - \varphi_H)} + (1 - \alpha)\Delta H_{\mathrm{Dis}}^0 \qquad (11)$$

This is the fundamental equation for the theoretical calculation of the heat of dilution of sulfuric acid. The data needed are the hypothetical heat of mixing, the degree of dissociation of the bisulfate ion, the standard heat of dissociation of the ion, and the heats of dilution of the hypothetical solutes $\mathrm{H \cdot H \cdot SO_4}$ and $\mathrm{H \cdot HSO_4}$.

The various terms were evaluated as follows: The hypothetical solutions of $\mathrm{H \cdot H \cdot SO_4}$ and $\mathrm{H \cdot HSO_4}$ formed by an imaginary separation of $\mathrm{H_2SO_4}$ into two parts were presumed to have the ionic strength of the original solution. Hence ΔH_M was taken as zero. For the next term we need the heat of dilution of $\mathrm{H \cdot H \cdot SO_4}$ from a molar concentration equal to one-third of the ionic strength to infinite dilution. The third term is $(1 - \alpha)$ times the heat of dilution of $\mathrm{H \cdot HSO_4}$ from a molarity equal to the ionic strength to infinite dilution. These quantities were computed from the equations for lithium sulfate and sodium chloride given by Young and Seligmann (64). Fortunately, the contribution of these terms is overshadowed by the last, and it is not likely that the error introduced by the substitution of these salts for the hypothetical acids increases appreciably the total error in the calculations.[6] The errors introduced by these and other approximations become smaller as concentration is decreased. The extrapolation aspect of these calculations is probably not affected at all by such approximations.

The standard heat of dissociation was determined from an empirical equation fitted to the equilibrium data obtained for temperatures between 5° and 55°C. by Young, Klotz, and Singleterry (57, 61). The equation gives for 25°C.: $\Delta F^0 = 2721$ cal. mole^{-1}, $\Delta H^0 = -5188$ cal. mole^{-1}, and $\Delta C_p^0 = -46$ cal. mole^{-1} degree.$^{-1}$

The results of the calculations are compared in figure 8 with the experimental data of Lange, Monheim, and Robinson (28). Their experimental values were obtained from a summation of a series of dilutions, some of which required precision in temperature measurement of 10^{-6}°C. or better. The authors' estimate of the probable error of a single point at the highest dilutions is 40–100 cal. mole.$^{-1}$ At least one of the two points nearest $C = 0$ must be in error by 100 cal. mole^{-1} or more, since it is clear that the original dilution data are internally inconsistent.

---

[6] Existing measurements of the heats of dilution and heat capacities indicate that acids and lithium salts are similar thermochemically (12, 51, 52). However, substitution of sodium sulfate and sodium chloride leads to little change in the results.

The agreement between the theoretical and experimental results over the full range of the data is wholly satisfactory in view of the limited data now available. The theoretical values become more accurate as concentration is reduced, whereas the experimental values become less accurate. Extrapolation based upon the theoretical curve should be very reliable. Since Lange, Monheim, and Robinson extrapolated linearly, all of their results for the integral heats of dilution (table 2 of reference 28) must be reduced by 290 cal. mole.$^{-1}$ Even from the data now available, this correction has been calculated with an uncertainty probably no greater than 10 or possibly 20 cal. mole.$^{-1}$ Better calculations of the curve near $\sqrt{m} = 0.2$ will have to await a complete experimental determination of the heats of dilution of an acid which is nearly completely dissociated, e.g., hydrochloric acid, further knowledge concerning heats of mixing and, especially, better knowledge of the degree of dissociation of the bisulfate ion.

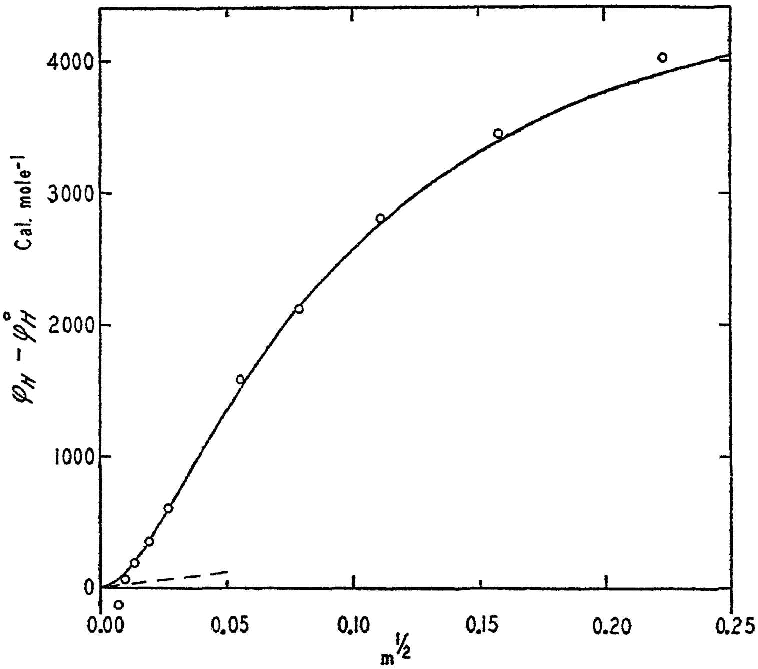


FIG. 8. Relative heat content of sulfuric acid. Curve is calculated. Points are experimental values corrected for error of linear extrapolation. Dashed line shows Debye–Hückel limiting law.

B. Apparent molal volume

The apparent molal volume of sulfuric acid decreases slowly as its concentration is decreased from 3 to about 0.3 molar (see figure 9). It then drops rapidly as the concentration is further decreased. From Klotz and Eckert (23) has come a quantitative explanation of these variations, in terms of the dissociation of the bisulfate ion and the relatively small variations in volume typical of electrolytes which are much more highly dissociated. The latter are ascribed largely to the interionic attractive forces treated by Debye and Hückel (*cf.* Redlich and Bigeleisen (45)). Klotz and Eckert employed the equation:

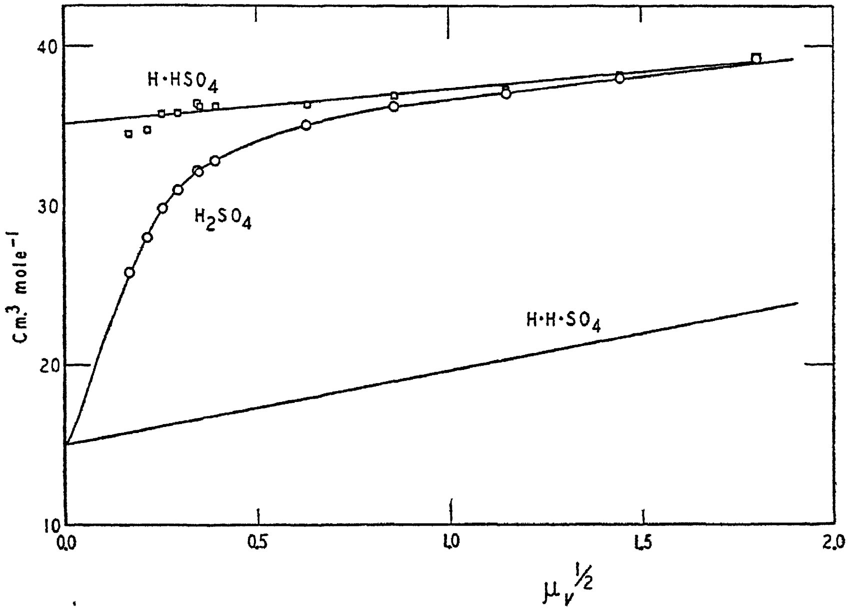


FIG. 9. Apparent molal volumes of H·HSO$_4$, H$_2$SO$_4$, and H·H·SO$_4$. Points are experimental.

$$\varphi_V(H_2SO_4) = \alpha\varphi_V(H\cdot H\cdot SO_4) + (1 - \alpha)\varphi_V(H\cdot HSO_4) \qquad (12)$$

and assumed that φ_V of each of the hypothetical solutes is determined (at constant temperature and pressure) by the total ionic strength. They calculated apparent molal volumes for the hypothetical solute H·H·SO$_4$, assuming that apparent molal volumes of ions are additive, and that each φ_V is determined by the total ionic strength.[5] For each of a series of ionic strengths they employed the equation:

$$\varphi_V(H\cdot H\cdot SO_4) = \varphi_V(K_2SO_4) + 2\varphi_V(HCl) - 2\varphi_V(KCl) \qquad (13)$$

The apparent molal volumes were determined from the respective density data for solutions of potassium sulfate, hydrochloric acid, and potassium chloride, all having ionic strengths equal to that of the sulfuric acid.

The lowest curve of figure 9 shows their results. The solid circles of figure 9 represent experimental determinations of the apparent molal volumes of sulfuric acid calculated by Klotz and Eckert from densities which they had determined with a precision of about one in 10^6. The open circles of the figure represent apparent molal volumes of $H\cdot HSO_4$ calculated by means of equation 12 from measured values of φ_V for H_2SO_4, the theoretical curve for $H\cdot H\cdot SO_4$, and the degrees of dissociation of HSO_4^-. The degrees of dissociation they estimated by a method which gave essentially the same results for the apparent molal volume as the upper γ_R line in figure 7. The fact that the curve so obtained for $H\cdot HSO_4$ is nearly linear within the experimental range and has about the slope expected for an acid of 1-1 valence type is evidence for the essential correctness of the procedures used.

Klotz and Eckert completed their beautiful analysis by estimating $\Delta\varphi_V^0$ for the dissociation of HSO_4^- (-20.2 cm.^{3} mole^{-1}) and calculating the portion of the curve for H_2SO_4 between the origin and the first experimental point.

C. Surface tension

Surface tensions of aqueous solutions of most strong electrolytes other than acids are nearly linear functions of concentration (20). In all of the measured surface tension isotherms of sulfuric acid there are prominent maxima. In the isotherm for 0°C. there is also a shallow minimum (see figure 10). The minimum becomes less pronounced as temperature rises and seems to disappear at about 55°C. (20, 34, 50).

Young and Grinstead (60) have recently attempted to determine to what extent varying degrees of dissociation are responsible for these phenomena. Observing that contributions of various ions to the surface tension are approximately additive except when hydrogen ion is in the presence of a second positive ion, they determined surface tension curves for dilute solutions of the hypothetical solute $H\cdot H\cdot SO_4$. The surface tension at any molality m' was taken to be the sum of the surface tension of a sodium sulfate solution of that molality plus the surface tension of a hydrochloric acid solution of twice that molality minus the surface tension of a sodium chloride solution of molality equal to $2\ m'$. The curves so computed are shown as short lines in figure 10. By a somewhat similar procedure curves were also determined for the other hypothetical solute $H\cdot HSO_4$.

The surface tension of sulfuric acid of molality m' was then computed as the sum of the surface tension of $H\cdot H\cdot SO_4$ at a molality $\alpha m'$ and the surface tension of $H\cdot HSO_4$ at a molality $(1 - \alpha)m'$.

Surface tension data for solutions of all electrolytes except hydrochloric acid were taken from the compilation by Young and Harkins in the *International Critical Tables* (20). For hydrochloric acid the tables were not adequate because they do not show temperature variations of $\Delta\gamma$, the difference between the surface tension of a solution and that of pure water. An early attempt to explain the curves of figure 10 met with complete failure because $\Delta\gamma$ of hydrochloric acid was assumed to be nearly independent of temperature, as it is for most strong electrolytes. Hoping to locate the cause of the failure, Young, Runge, and Grinstead (53, 63) made measurements at 0°, 20°, and 40°C. They found the surface tension lowering of dilute hydrochloric acid to vary rapidly with temperature; the lowering is much larger at 0°C. than at 18°C. At 40°C. $\Delta\gamma$ is actually positive for all concentrations between 0 and 8 molal.

The new surface tension data led to very different theoretical calculations. Table 4 contains a comparison of experimental and theoretical points on the 0°C. and 18°C. isotherms. The agreement is nearly as good as might be expected if there were no uncertainty in the degrees of dissociation estimated from a single approximate γ_R curve for both 0°C. and 18°C.

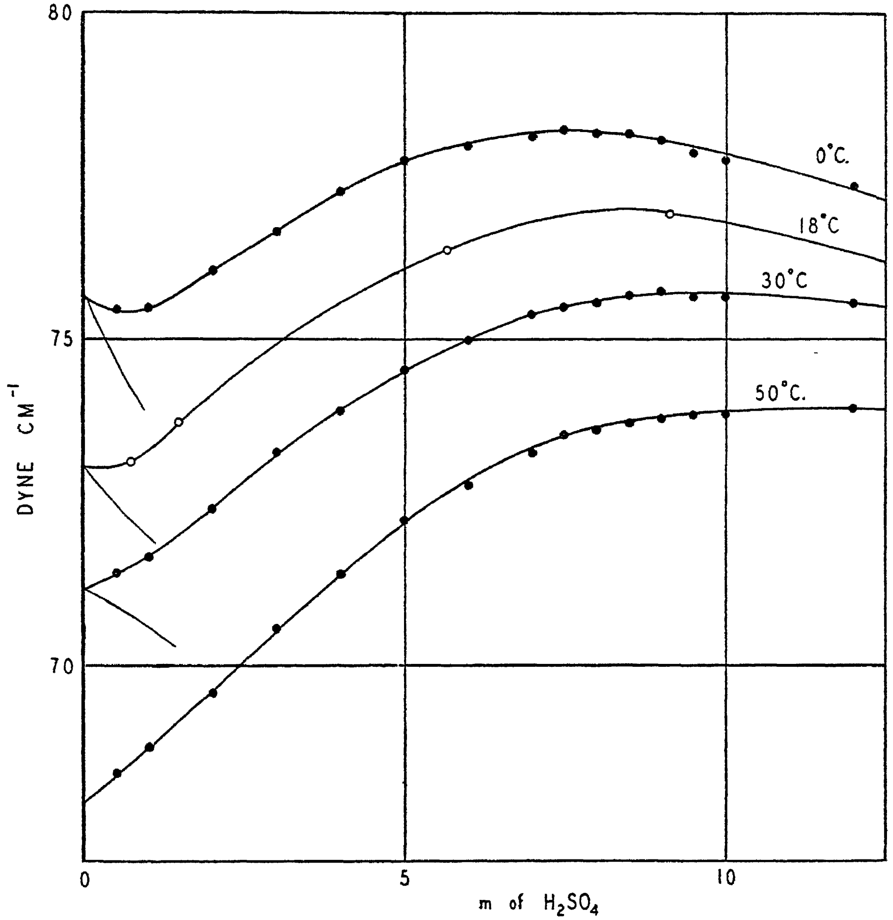


FIG. 10. Surface tension (dynes cm.$^{-1}$) of sulfuric acid solutions at 0°, 18°, 30°, and 50°C. Light lines are theoretical curves for H·H·SO_4.

Estimation of the height and position of the maximum surface tension is more difficult. Some of the principles and approximations acceptable when the concentrations are small become unreliable at higher concentrations. Probably the most serious source of error is the neglect, in this simple procedure, of the effects of ·$HHSO_4$. According to the calculations made without any allowance for the effects of ·$HHSO_4$ the maximum $\Delta\gamma$ should occur in the solution of about the concentration in which the hydrogen-ion and bisulfate-ion concentrations are maximal (65 weight per cent). Actually, $\Delta\gamma$ reaches a maximum of 3.9 dynes cm.$^{-1}$ in a 47 per cent solution.

A rough estimate of the effect of $\cdot HHSO_4$ in the 47 per cent solution may be made. Consider the solution to be one containing the solute ions H^+, HSO_4^-, and SO_4^{--} in a mixed solvent consisting of H_2O and $\cdot HHSO_4$. The mole fraction of $\cdot HHSO_4$ in the hypothetical solvent (separated from the solute ions) would be 0.024. The surface tension of the solvent would probably be less than that of pure water. Linear interpolation on a mole fraction scale between γ of water and γ of pure sulfuric acid leads to a value for γ for the mixture which is 0.5 dyne cm.$^{-1}$ less than that of water. Surface tensions of mixtures are usually considerably lower than linear interpolation would indicate. The correction to be expected, therefore, is larger than 0.5, perhaps as much as twice that. The difference between the experimental maximum $\Delta\gamma$ (3.9), and $\Delta\gamma$ calculated (5.0) for the 47 per cent solution is therefore about what is to be expected.

TABLE 4

Comparison of calculated and experimental points on surface tension isotherms

	0°C.		18°C.	
	Calculated	Observed	Calculated	Observed
m at which $\Delta\gamma = 0$	1.5	1.3	0.04	Small*
m at which $d\gamma/dm = 0$	0.65	0.5–0.7	0.02	Small*
Minimum $\Delta\gamma$	−0.15	−0.21	−0.01	Small*
$\Delta\gamma$ at $m = 0.7$	−0.15	−0.21	−0.25	0.06
$\Delta\gamma$ at $m = 1.5$			0.72	0.67

* Too small to estimate.

The surface tension results constitute a strong argument for the correctness of the picture of sulfuric acid coming from the Raman data. There is additional support for the picture in conductance data. The product of viscosity and specific conductance is greatest in a solution of about the same composition as that in which the hydrogen-ion concentration is largest (25).

D. Apparent molal heat capacity

The equation for the theoretical calculation of the relative apparent molal heat capacity is obtained from equation 11 by differentiation of both sides with respect to temperature. When ΔH of mixing is negligible, the equation is:

$$\underset{H_2SO_4}{\varphi_C - \overset{0}{\varphi_C}} = [\Delta H^0_{\text{Dis}} + \underset{H\cdot H\cdot SO_4}{(\varphi_H - \overset{0}{\varphi_H})} - \underset{H\cdot HSO_4}{(\varphi_H - \overset{0}{\varphi_H})}]\left(\frac{\partial\alpha}{\partial T}\right)_m$$

$$+ (\alpha - 1)\Delta C^0_{\text{Dis}} + \alpha\underset{H\cdot H\cdot SO_4}{(\varphi_C - \overset{0}{\varphi_C})} + (1 - \alpha)\underset{H\cdot HSO_4}{(\varphi_C - \overset{0}{\varphi_C})} \quad (14)$$

$$+ \alpha \underset{H\cdot H\cdot SO_4}{\left(\frac{\partial\varphi_H}{\partial\mu}\right)_T\left(\frac{\partial\mu}{\partial T}\right)_m} + (1 - \alpha)\underset{H\cdot HSO_4}{\left(\frac{\partial\varphi_H}{\partial\mu}\right)_T\left(\frac{\partial\mu}{\partial T}\right)_m}$$

In equation 14 φ_H and φ_C denote properties of the hypothetical solutes in solutions of ionic strengths equal to the ionic strength of the sulfuric acid solution.

There have been two excellent experimental studies of the apparent molal heat capacities of sulfuric acid: one by Biron (3) in 1899 and a more recent one by Randall and Taylor (38). Their determinations of φ_C are in excellent agreement. $\varphi_C - \overset{0}{\varphi}_C$, the relative apparent molal heat capacity, was determined by Randall and Taylor only. For that purpose they computed the apparent molal heat capacity of infinitely dilute sulfuric acid from available published values of $\overset{0}{\varphi}_C$ for hydrochloric acid, sodium sulfate, and sodium chloride. The limiting heat capacities had been determined by linear extrapolation to infinite dilution of measured values of φ_C of the respective electrolytes. This procedure has been shown by Young and Machin (62) to be erroneous. In general, it leads to values of $\overset{0}{\varphi}_C$ which are too low.

Data are not now available for a satisfactory demonstration of the use of equation 14. Nevertheless, we have ventured to make a calculation based upon the γ_R curves of figure 7, and the equilibrium constants determined by Young, Klotz, and Singleterry (57, 61). In the absence of any knowledge concerning the variation of γ_R with temperature it was necessary that its temperature coefficient be treated as though it were zero. This approximation is especially unsatisfactory for the calculation of a temperature derivative. Thermal data for the solutions of lithium salts used as analogues of $H \cdot H \cdot SO_4$ and $H \cdot HSO_4$ were also grossly inadequate. When necessary, values of relative apparent molal heat contents and heat capacities were estimated from the tables of properties of similar electrolytes compiled by Gucker and Schminke (12), Young and Seligmann (64), Randall and Rossini (37), and Rossini (51, 52). The result is shown in table 5. The agreement is entirely satisfactory; the discrepancy is no larger than is to be expected as a result of the various approximations. It is not much larger than the uncertainty (due to extrapolation) in the experimental value itself.

TABLE 5

Apparent molal heat capacity of sulfuric acid

C	EXPERIMENTAL		CALCULATED
	φ_C	$\varphi_C - \varphi_C^0$	$\varphi_C - \varphi_C^0$
0.1	9	76	65

The discrepancy in this calculation, like those in others discussed above, points to the need for adequate optical, thermodynamic, and other relevant data. The demand is rather for measurements made over sufficiently wide ranges of temperature, concentration, and pressure than merely for improved accuracy. When the necessary information is available, it should be possible to increase the beauty and precision of calculations such as these and greatly improve our understanding of the behavior of solutions of those interesting electrolytes which are moderately strong but by no means completely dissociated.

We are grateful to Professor G. Herzberg for an inspiring conversation concerning these calculations. Our hearty thanks are also extended to Messrs. L. F. Maranville and H. M. Smith and to Dr. C. R. Singleterry, who have helped with calculations and handling of the data.

REFERENCES

(1) BELL, R. M., AND JEPPESEN, M. A.: J. Chem. Phys. **2**, 711 (1934).
(2) BELL, R. M., AND JEPPESEN, M. A.: J. Chem. Phys. **3**, 245 (1935).
(3) BIRON, E. VON: J. Russ. Phys. Chem. Soc. **31**, 171 (1899).
(4) BRAY, W. C., AND LIEBHAFSKY, H. A.: J. Am. Chem. Soc. **57**, 51 (1935).
(5) BRØNSTED, J. N.: Z. phyzik. Chem. **68**, 693 (1910).
(6) CHÉDIN, J.: Ann. chim. [11] **8**, 243 (1937).
(7) DALMON, R.: Compt. rend. **207**, 473 (1938).
(8) DE BÉTHUNE, A. J., AND KIMBALL, G. E.: J. Chem. Phys. **13**, 53 (1945).
(9) DEBYE, P., AND HÜCKEL, E.: Physik. Z. **24**, 185 (1923).
(10) GERDING, H., AND NIJVELD, W. J.: Rec. trav. chim. **59**, 1206 (1940).
(11) GERDING, H., NIJVELD, W. J., AND MULLER, G. J.: Nature **137**, 1033 (1936).
(12) GUCKER, F. T., JR., AND SCHMINKE, K. H.: J. Am. Chem. Soc. **54**, 1358 (1932).
(13) HALBAN, H. VON, AND BRÜLL, J.: Helv. Chim. Acta **27**, 1719 (1944).
(14) HALBAN, H. VON, AND EISENBRAND, J.: Z. physik. Chem. **132**, 401 (1928).
(15) HALBAN, H. VON, AND KÖRTUM, G.: Z. physik Chem. **A170**, 351 (1934).
(16) HAMMETT, L. P., AND DEYRUP, A. J.: J. Am. Chem. Soc. **54**, 2721 (1932).
(17) HARNED, H. S., AND OWEN, B. B.: *The Physical Chemistry of Electrolytic Solutions.* Reinhold Publishing Corporation, New York (1943).
(18) HERZBERG, G.: *Molecular Spectra and Molecular Structure. II. Infrared and Raman Spectra of Polyatomic Molecules*, p. 178. D. Van Nostrand Company, Inc., New York (1945).
(19) HIBBEN, J. H.: *The Raman Effect and Its Chemical Applications.* Reinhold Publishing Corporation, New York (1939).
(20) *International Critical Tables of Numerical Data; Physics, Chemistry and Technology.* McGraw-Hill Book Company, Inc., New York (1926–33).
(21) KIELLAND, J.: J. Am. Chem. Soc. **59**, 1675 (1937).
(22) KLOTZ, I. M.: Thesis, University of Chicago, 1940.
(23) KLOTZ, I. M., AND ECKERT, C. F.: J. Am. Chem. Soc. **64**, 1878 (1942).
(24) KNIETSCH, R.: Ber. **34**, 4069 (1901).
(25) KOHLRAUSCH, F.: Ann. Physik [2] **159**, 233 (1876).
(26) KORTÜM, G., AND HALBAN, H. VON: Z. physik. Chem. **A170**, 212 (1934).
(27) KOSSIAKOFF, A., AND HARKER, D.: J. Am. Chem. Soc. **60**, 2047 (1938).
(28) LANGE, E., MONHEIM, J., AND ROBINSON, A. L.: J. Am. Chem. Soc. **55**, 4733 (1933).
(29) LEWIS, G. N., AND RANDALL, M.: *Thermodynamics and the Free Energy of Chemical Substances.* McGraw-Hill Book Company, Inc., New York (1923).
(30) MACINNES, D. A.: J. Am. Chem. Soc. **48**, 2068 (1926).
(31) MACINNES, D. A.: *The Principles of Electrochemistry.* Reinhold Publishing Corporation, New York (1939).
(32) MACINNES, D. A., AND SHEDLOVSKY, T.: J. Am. Chem. Soc. **54**, 1429 (1932).
(33) MACINNES, D. A., SHEDLOVSKY, T., AND LONGSWORTH, L. G.: Chem. Revs. **13**, 29 (1933).
(34) MORGAN, J. L. R., AND DAVIS, C. E.: J. Am. Chem. Soc. **38**, 555 (1916).
(35) NISI, H.: Japan. J. Phys. **5**, 119 (1929).
(36) PAI, N. G.: Phil. Mag. **20**, 616 (1935).
(37) RANDALL, M., AND ROSSINI, F. D.: J. Am. Chem. Soc. **51**, 323 (1929).
(38) RANDALL, M., AND TAYLOR, M. D.: J. Phys. Chem. **45**, 959 (1941).
(39) RAO, I. R.: Proc. Roy. Soc. (London) **A144**, 159 (1934).

(40) Rao, N. R.: Indian J. Phys. **14**, 143 (1940).
(41) Rao, N. R.: Indian J. Phys. **15**, 185 (1941).
(42) Rao, N. R.: Indian J. Phys. **16**, 71 (1942).
(43) Rao, N. R.: Indian J. Phys. **17**, 332 (1943).
(44) Redlich, O.: Chem. Revs. **39**, 333 (1946).
(45) Redlich, O., and Bigeleisen, J.: Chem. Revs. **30**, 171 (1942).
(46) Redlich, O., and Bigeleisen, J.: J. Am. Chem. Soc. **65**, 1883 (1943).
(47) Redlich, O., Holt, E. K., and Bigeleisen, J.: J. Am. Chem. Soc. **66**, 13 (1944).
(48) Redlich, O., and Nielsen, L. E.: J. Am. Chem. Soc. **65**, 654 (1943).
(49) Rogan, Ruth: Thesis, University of Chicago, 1948.
(50) Röntgen, W. C., and Schneider, J.: Ann. Physik [3] **29**, 165 (1886).
(51) Rossini, F. D.: Bur. Standards J. Research **6**, 791 (1931).
(52) Rossini, F. D.: Bur. Standards J. Research **7**, 47 (1931).
(53) Runge, P. W.: Thesis, University of Chicago, 1947.
(54) Scheibe, G.: Ber. **59**, 2616 (1926).
(55) Sherrill, M. S., and Noyes, A. A.: J. Am. Chem. Soc. **48**, 1861 (1926).
(56) Simons, L.: Soc. Sci. Fennica Commentationes Phys. Math. **7**, No. 9 (1934).
(57) Singleterry, C. R.: Thesis, University of Chicago, 1940.
(58) Smith, M. B.: Thesis, University of Chicago, 1942.
(59) Woodward, L. A., and Horner, R. G.: Proc. Roy. Soc. (London) **144A**, 129 (1934).
(60) Young, T. F., and Grinstead, S. R.: Ann. N. Y. Acad. Sci., in press.
(61) Young, T. F., Klotz, I. M., and Singleterry, C. R.: Unpublished.
(62) Young, T. F., and Machin, J. S.: J. Am. Chem. Soc. **58**, 2254 (1936).
(63) Young, T. F., Runge, P. W., and Grinstead, S. R.: Unpublished.
(64) Young, T. F., and Seligmann, P.: J. Am. Chem. Soc. **60**, 2379 (1938).

ACTIVITY COEFFICIENTS AND DISSOCIATION OF LEAD CHLORIDE IN AQUEOUS SOLUTIONS[1]

ROBERT M. GARRELS AND FRANK T. GUCKER, JR.

Department of Geology, Northwestern University, Evanston, Illinois, and Departments of Chemistry, Northwestern University, Evanston, Illinois, and Indiana University, Bloomington, Indiana

Received October 25, 1948

The study of lead chloride is taken as an example of the experimental techniques and theoretical methods which may be applied to all moderately strong electrolytes. The deviation of the behavior of lead chloride from that of a strong electrolyte has been explained on the basis of the Gronwall–LaMer–Sandved extension of the Debye–Hückel theory for small highly charged ions, and also on the basis of incomplete dissociation. The second interpretation is strengthened by the discovery of ultraviolet absorption bands identified with the $PbCl^+$ ion, a quantitative investigation of which affords one method of determining the dissociation constant of this ion. Study of the conductance of lead chloride solutions also gives evidence of incomplete dissociation and allows an evaluation of the dissociation constant. Finally, a great deal of information can be obtained from the E.M.F. of cells containing lead chloride alone and in the presence of other electrolytes, from which its activity coefficients have been determined. We present some new experimental data for the E.M.F. of cells with added sodium chloride, over a wide range of temperature and concentration. The plot of the usual extrapolation function $E^{0\prime}$ against the square root of the stoichiometric ionic strength approaches the Debye–Hückel limiting law from above rather than below, thus making extrapolation to find E^0 difficult. The results with added barium nitrate lie along a different curve from those with added alkali chlorides, showing that ionic strength is not the only variable to be considered. Nearly all of the E.M.F. data may be correlated, however, if we assume a secondary dissociation constant of about 0.03, calculate the ionic concentrations, and plot a suitably defined function against the actual ionic strength rather than the stoichiometric ionic strength. The same dissociation constant fits the conductance data, but a considerably larger value of the constant is obtained from the absorption spectra, which emphasizes the difficulties often involved in dealing with moderately strong electrolytes.

INTRODUCTION

There have been several studies of the activities of lead chloride in aqueous solution. In 1906 J. N. Brönsted (5) made a series of measurements from 0° to 90°C., using the cell:

$$\text{Pb(0.72 per cent amalgam)} \mid PbCl_2(m) \mid AgCl(s)\text{–}Ag(s) \qquad (1)$$

Subsequent workers have found that the silver–silver chloride electrodes of the type devised by Jahn, which Brönsted used, are not adequately reproducible, and that the lead amalgam electrode is only stable between 1.4 and 66 per cent lead.

---

[1] Presented at the Symposium on Thermodynamics and Molecular Structure of Solutions which was held under the auspices of the Division of Physical and Inorganic Chemistry at the 114th Meeting of the American Chemical Society, Portland, Oregon, September 13 and 14, 1948.

Later, A. J. Allmand and E. Hunter (3) measured the same cell at 25°C. over a wide range of concentration. They used amalgamated lead sticks for the lead electrodes, and carefully excluded oxygen from the system. They found a value of $E^0 = 0.3396$ v. for the cell, and obtained activity coefficients for lead chloride which agreed fairly well with those of Brönsted.

W. R. Carmody (6) carried out a careful set of measurements of the cell:

$$\text{Pb(two-phase amalgam)} \mid \text{PbCl}_2(m) \mid \text{AgCl}(s)\text{–Ag}(s) \qquad (2)$$

He also worked at 25°C. and varied the lead chloride concentration from saturation (0.039 m) to 0.0002 m. He found a value of $E^0 = 0.3426$ v. for the cell by linear extrapolation.

In 1936 M. C. Hannan (14) studied the same cell at concentrations from 0.01 to 0.0005 m, at 12.5° intervals from 0° to 50°C. She used two-phase lead amalgam electrodes prepared by electrolysis with a mercury cathode, silver–silver chloride electrodes of the type used by Carmody, and an atmosphere of nitrogen in the cell. Her work was done to test the Debye–Hückel theory as extended to unsymmetrical electrolytes by LaMer, Gronwall, and Grieff (19). Using their extended equation she found a value of $E^0 = 0.3435$ at 25°C.

A number of studies have been made of the activities of lead chloride in the presence of other electrolytes, at 25°C. Most of these (1, 2, 4, 20, 18, 17), dealing with the behavior of lead chloride in the presence of lithium, sodium, or potassium chloride, were carried out in connection with phase-rule studies of these systems. The last article, by E. R. Hounsell and H. N. Parton (17) summarizes the work on E.M.F. and activity coefficients, based on a value of $E^0 = 0.3432$ v. The authors point out that much of the previous work in the series was vitiated by dissolved oxygen in the solutions dilute in lead chloride. They used an atmosphere of hydrogen instead of nitrogen, and found that duplicate cells remained constant for 12 hr. or more.

E. Güntelberg (13) also measured the E.M.F. of cells containing lead chloride alone and in the presence of hydrochloric acid and potassium chloride. Unfortunately his work is available to us only in abstract, and we do not have the details of experimentation or his experimental results. His conclusions will be discussed in a later section.

H. D. Crockford and H. O. Farr, Jr., (7) studied the E.M.F. of cells containing lead chloride in the presence of barium nitrate, in order to determine the value of the a parameter (distance of closest approach of the ions) in mixed electrolytes.

Our work extends the measurements on lead chloride in the presence of sodium chloride from 25° to 65°C. The E.M.F. data are used to calculate the activity coefficient of lead chloride in these solutions. Its abnormally low value in these solutions is interpreted in terms of ionic association.

Elsewhere, one of us (12) has discussed the geological implications of these results, in connection with the interesting geological problem which started us upon this work and which was posed by the fact that, in vertical vein deposits, galena (PbS) frequently is found farther from the place of origin than sphalerite (ZnS), despite the lower solubility of the lead sulfide in most solutions. Since primary liquid inclusions in galena crystals contain high concentrations (up to 4 *m*) of soluble chlorides, we investigated the effect of added sodium chloride upon the activity coefficient of lead chloride over the temperature range of geologic interest. The results were used to explain the increased solubility of lead sulfide in relation to zinc sulfide in vein solutions.

EXPERIMENTAL METHODS AND RESULTS

We used materials similar to those which Hounsell and Parton (17) found satisfactory except that tank nitrogen (Linde) instead of hydrogen was used to provide an oxygen-free atmosphere, since none of our solutions was extremely dilute in lead chloride. This nitrogen was purified by passing it through a series of gas-washing bottles containing successively potassium permanganate, alkaline pyrogallol, and cuprous nitrate, and then over copper turnings heated to 300–400°C.

TABLE 1

E.M.F. of cell: Pb *(two-phase amalgam)* | $PbCl_2(m_1)$, $NaCl(m_2)$ | $AgCl(s)$-$Ag(s)$, *from 25° to 65°C.; corresponding values of* E^0 *and mean activity coefficients*

m_1 ($PbCl_2$)	m_2 (NaCl)	$\mu = 3m_1 + m_2$		25°C.	35°C.	45°C.	55°C.	65°C.
0.000	0.000	0.000	E^0	0.3434	0.3404	0.3368	0.3229	0.3287
0.0018	0.020	0.0254	E..	0.5418	0.5447	0.5474	0.5502	0.5526
			$\gamma_\pm$....	0.581	0.591	0.597	0.595	0.595
0.0036	0.100	0.1108	E........ .	0.5054	0.5088	0.5114	0.5140	0.5164
			$\gamma_\pm$	0.432	0.422	0.414	0.404	0.394
0.0018	0.200	0.2054	E....... ...	0.5075	0.5124	0.5161	0.5195	0.5221
			$\gamma_\pm$....... . .	0.336	0.316	0.304	0.291	0.284
0.0018	0.400	0.4054	E..	0.5025	0.5067	0.5099	0.5130	0.5161
			$\gamma_\pm$..........	0.243	0.231	0.224	0.215	0.206
0.0018	0.600	0.6054	E..........	0.5004	0.5038	0.5073	0.5103	0.5136
			$\gamma_\pm$..........	0.196	0.190	0.182	0.176	0.167

The cell was designed according to principles used by Carmody (6), with the addition of a second cell unit so that the E.M.F. readings could be checked. Cell operation followed Carmody's procedure.

Lead amalgam electrodes were prepared within the cell to avoid oxidation. The air in the cells was displaced with nitrogen; then the mercury and lead successively were dropped in against a strong current of nitrogen, and the amalgam was heated to 80°C. to insure saturation at all experimental temperatures.

Silver chloride electrodes were prepared according to the directions of M. Dole (9). They were white when fresh and gradually turned brown after use. They were reproducible within 0.1 mv., and the change in color did not affect the potential.

Temperature was controlled within 0.1°C. by immersion in an oil bath heated by a resistance coil and controlled by a mercury thermoregulator. Closer control was not necessary because of the small temperature coefficient of the cell.

E.M.F. readings were taken to 0.1 mv. on a Type K potentiometer, made by Leeds and Northrup Company, balanced against a recently calibrated standard cell. The experimental E.M.F. values of the two cells always checked within 0.6 mv., and usually within 0.2 mv.

Our experimental results are given in table 1.

E^0 AND MEAN ACTIVITY COEFFICIENTS

In cells of the type discussed here, either in pure lead chloride solutions or in mixtures of lead chloride and other salts, the reaction is:

$$\mathrm{Pb(two\text{-}phase\ amalgam)} + 2\mathrm{AgCl}(s) = 2\mathrm{Ag}(s) + \mathrm{PbCl_2}(m)$$

The E.M.F. of this cell E in terms of the standard potential E^0 and the activity a of lead chloride is given by the equation:

$$E = E^0 - \frac{RT}{nF} \ln a_{\mathrm{PbCl_2}} \tag{1}$$

or

$$E = E^0 - \tfrac{1}{2}k \log a_{\mathrm{PbCl_2}} \tag{2}$$

where $k = 2.3029\ RT/F$. This also can be written:

$$E = E^0 - \tfrac{1}{2}k \log m_{\mathrm{Pb^{++}}} \cdot m^2_{\mathrm{Cl^-}} \cdot \gamma_{\mathrm{Pb^{++}}} \cdot \gamma^2_{\mathrm{Cl^-}} \tag{3}$$

In terms of the mean activity coefficient of the lead chloride, the equation may be rearranged as follows:

$$E + \tfrac{1}{3}k' \log m_1(2m_1 + m_2)^2 = E^0 - k' \log \gamma_\pm \tag{4}$$

where m_1 is the concentration of lead chloride, m_2 is the concentration of added chloride, and $k' = \frac{3}{2}k$.

The left-hand member of equation 4, usually designated $E^{0\prime}$, is used in extrapolation to find E^0 and $\log \gamma_\pm$. In solutions of lead chloride alone, this equation becomes:

$$E + \tfrac{1}{3}k' \log 4 + k' \log m_1 = E^0 - k' \log \gamma_\pm \tag{5}$$

We have used equations 4 and 5 to calculate $E^{0\prime}$ from our results and those of the other investigators included in table 2. In the sixth column are given values of $E^{0\prime}/k'$ which in figures 1 and 2 are plotted as ordinates instead of $E^{0\prime}$ in order to separate the curves for the different temperatures. The abscissas are $\mu^{1/2}$. Here μ is the usual ionic strength, defined as:

$$\mu \equiv \frac{1}{2} \sum_i m_i z_i^2 \tag{6}$$

where m_i is the stoichiometric molality and z_i the charge of the i^{th} ion. We have included on the 25°C. curve the results of Carmody (6) and of Hannan (14) for cells containing lead chloride alone, and representative values of Hounsel and Parton (17) for cells with added lithium, sodium, or potassium chloride. The results show that, in all these solutions, the value of $E^{0\prime}$ depends chiefly upon the ionic strength. However there are some deviations, particularly for the solutions with a high concentration of added potassium chloride.

TABLE 2

E.M.F. of cell: Pb (*two-phase amalgam*) | $PbCl_2(m_1)$, $MCl(m_2)$, $Ba(NO_3)_2(m_3)$ | AgCl(*s*)–Ag(*s*) *at 25°C., with various extrapolation functions**

$m_1(PbCl_2)$	$m_2(MCl)$	$m_3(Ba(NO_3)_2)$	μ	E(EXPTL.)	$E^{0\prime}/k'$	$E^{0\prime\prime}/k'$	$E^{0\prime\prime\prime}/k'$	μ'
0.0002116			0.000635	0.6537	3.893	3.867	3.864	0.000629
0.0006197			0.001859	0.6143	3.916	3.872	3.863	0.001817
0.001034			0.003102	0.5960	3.932	3.876	3.861	0.002994
0.001337			0.004011	0.5870	3.942	3.878	3.861	0.003839
0.002348			0.007044	0.5677	3.969	3.884	3.857	0.006574
0.002620			0.007860	0.5639	3.974	3.884	3.855	0.007293
0.005160			0.01548	0.5419	4.020	3.894	3.848	0.01370
0.0018	0.0200		0.0254	0.5418	4.106	3.945	3.887	0.02427
0.01039			0.03117	0.5205	4.083	3.904	3.837	0.02575
0.02048			0.06144	0.5012	4.160	3.909	3.819	0.04656
0.019037		0.00503	0.07220	0.50426	4.1632	3.891	3.805	0.05939
0.02955			0.08865	0.4913	4.208	3.907	3.804	0.06366
0.0036	0.100		0.1108	0.5054	4.235	3.898	3.771	0.1066
0.03905			0.11715	0.4842	4.249	3.903	3.793	0.08253
0.019037		0.03030	0.14801	0.50872	4.2135	3.824	3.745	0.1369
0.0018	0.171		0.1764	0.5089	4.315	3.890	3.731	0.1740
0.0018	0.200		0.2054	0.5075	4.344	3.886	3.715	0.2029
0.0072	0.201		0.2226	0.4896	4.359	3.882	3.710	0.2125
0.019037		0.06069	0.23918	0.51223	4.2530	3.758	3.684	0.2291
0.019037		0.08027	0.29792	0.51430	4.2764	3.724	3.653	0.2882
0.0018	0.342		0.3474	0.5032	4.448	3.852	3.642	0.3446
0.019037		0.10025	0.35786	0.51543	4.2891	3.684	3.615	0.3486
0.0018	0.400		0.4054	0.5025	4.485	3.841	3.619	0.4025
0.0072	0.469		0.4906	0.4851	4.542	3.833	3.600	0.4790
0.0018	0.600		0.6054	0.5004	4.578	3.791	3.535	0.6024
0.0009	0.697		0.6997	0.5110	4.640	3.793	3.523	0.6982
0.0072	0.697		0.7186	0.4867	4.672	3.815	3.547	0.7064

* Data of Carmody (6) for lead chloride alone; results of this research and of Hounsell and Parton (17) for added alkali halides, and of Crockford and Farr (7) for added barium nitrate. For definitions of $E^{0\prime}$, $E^{0\prime\prime}$, $E^{0\prime\prime\prime}$, μ, and μ' see, respectively, equations 4, 12, 31, 6, and 34.

Our measurements did not extend to low enough concentrations to fix the dilute end of the curve with certainty. For this purpose we used the results of Hannan, which covered the range from 0° to 50°C. She had calculated E^0 at each temperature by means of the Debye–Hückel equation as extended by LaMer, Gronwall, and Grieff (19). She found that the results over the whole range of temperature were represented by a single value of the parameter $\mathring{a}$ = 1.75 Å. Although this value is too small to have much physical significance, these results were the most reliable ones available over a range of temperature.

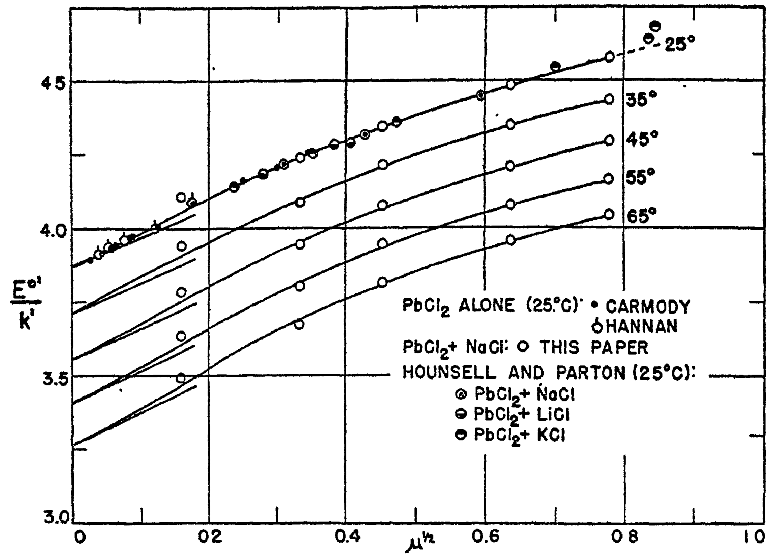


FIG. 1. $(E^{0\prime}/k')$ *vs.* $\mu^{1/2}$ for lead chloride alone and with added alkali halides at 25–65°C.

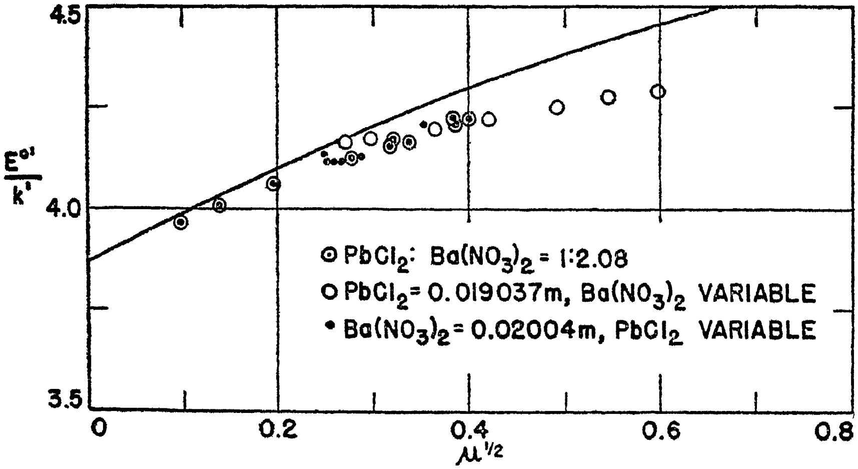


FIG. 2. $(E^{0\prime}/k')$ *vs.* $\mu^{1/2}$ for lead chloride with added barium nitrate at 25°C.

We found that the E^0 values she tabulated at 0°, 12.5°, 25°, 37.5°, and 50°C. could be represented within ±0.04 mv. by a slight modification of the equation she had proposed to relate E^0 to the Centigrade temperature, namely:

$$E^0 = 0.34790 - 6.77 \times 10^{-5}t - 4.990 \times 10^{-6}t^2 + 2.270 \times 10^{-8}t^3 \qquad (7)$$

This equation could be extended to 65°C. with no great loss of reliability. The

E^0 values given in table 1, calculated from this equation, probably are reliable to within about 0.5 mv. This limit is imposed by the discrepancies between different methods of extrapolation, which will be discussed later. The values of the mean activity coefficients listed in table 1 are calculated from E and E^0 by means of equation 4.

The Debye–Hückel limiting law slopes are shown at each temperature. It will be seen that the experimental curves do not approach these from below, as in the case of typical strong electrolytes. Instead, they lie above the limiting law until an ionic strength of about 0.7 is reached. The shape of the curve makes the usual extrapolation for E^0 uncertain.

Figure 1 does not include the results of Crockford and Farr (7) for lead chloride in the presence of barium nitrate solutions. When these are plotted on the same scale, they deviate widely from the results obtained in the presence of added chlorides. To avoid confusion of the experimental points, the results of Crockford and Farr are plotted separately in figure 2, together with the solid line of figure 1 at 25°C. Three series of experiments were carried out. In the first, the ratio of the molalities of lead chloride to barium nitrate was kept constant; in the second, the concentration of lead chloride was held constant and that of the barium nitrate was changed; while in the third the barium nitrate molality was fixed and that of the lead chloride was varied. The three series agree fairly well among themselves, but at high concentration they fall far below the results obtained in the presence of added chloride.

Crockford and Farr interpreted their results in terms of the change in the distance of closest approach of the ions in these mixtures. They used the equation derived by H. D. Crockford and H. C. Thomas (8) for the mean distance of closest approach $\bar{a}$ in a mixture of two salts:

$$\bar{a} = \frac{a_1 f_1 + a_2 f_2}{f_1 + f_2} \tag{8}$$

Here a_1 and a_2 are the ionic diameter parameters for the individual salts and f_1 and f_2 are the respective collision frequencies, assumed to be proportional to the squares of the concentrations of the two salts. They used a_1 = 1.75 Å. for lead chloride, calculated by LaMer, Gronwall, and Grieff (19), and a_2 = 2.68 Å. for barium nitrate. This was computed from the value of a = 2.50 Å., determined for the first series of experiments by the method of LaMer, Gronwall, and Grieff. This is appreciably less than might be expected. The crystal ionic radius given by L. Pauling (21) for Ba^{++} (1.35 Å.) is about that for K^+ (1.33 Å.) and G. Scatchard, S. S. Prentiss, and P. T. Jones (26), from their studies of freezing points of alkali nitrates, conclude that the nitrate ion is intermediate between the bromide and iodide ions in size in solution. The values for potassium bromide and iodide are given by H. S. Harned and B. B. Owen (16) based on calculations from Hückel's equation:

$$\log \gamma_\pm = -\frac{z_1 z_2 A\sqrt{\mu}}{1 + \mathring{a}B\sqrt{\mu}} + C\mu \tag{9}$$

where A and B are the constants of the Debye–Hückel theory and C is an empirical constant. Taking the mean of these values gives for barium nitrate the figure $\mathring{a} = 3.9$ Å. This is so near the values of 4.2, 4.0, and 3.8 given by Harned and Owen for lithium, sodium, and potassium chloride, respectively, that the strikingly different effect of barium nitrate upon lead chloride hardly can be due to the $\mathring{a}$ parameter.

THE ACTIVITY OF LEAD ION: COMPLEX FORMATION

As many previous workers have pointed out, the activity coefficients of lead chloride are unusually low in solutions of increasing ionic strength. Normal strong 2–1 electrolytes in very dilute solution obey the Debye–Hückel equation which, in terms of ionic strength, is:

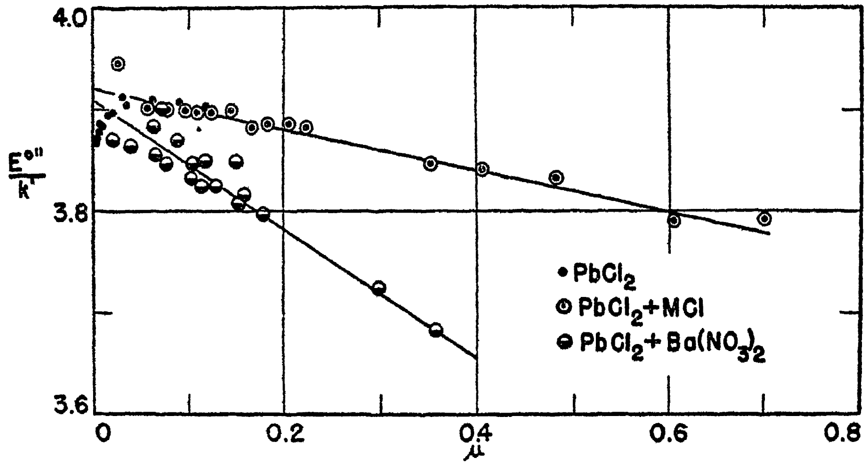


FIG. 3. $(E^{0''}/k')$ *vs.* μ for lead chloride alone and with added electrolytes at 25°C.

$$\log \gamma_{\pm} = -\frac{2A\sqrt{\mu}}{1 + \mathring{a}B\sqrt{\mu}} \tag{10}$$

This is sometimes divided out to give approximately:

$$\log \gamma_{\pm} = -2A\sqrt{\mu} + B'\mu \tag{11}$$

where $B' = 2AB\mathring{a}$. Substituting for $\log \gamma_{\pm}$ from equation 11 into equation 4 and rearranging terms gives:

$$E + \tfrac{1}{3}k' \log m_1(2m_1 + m_2)^2 - 2k'A\sqrt{\mu} = E^0 - k'B'\mu \tag{12}$$

The left-hand member of this equation, designated $E^{0''}$, should decrease nearly linearly with μ and frequently is used in extrapolating data to obtain E^0. In the case of lead chloride, however, a very different result is obtained, as shown in figure 3, which is a plot of the values of $E^{0''}/k'$ given in the seventh column of table 2. Our results and those of Hounsell and Parton on mixtures of lead chloride and alkali chlorides fall nearly together along a reasonable line, while the results of Crockford and Farr in mixtures of lead chloride and barium nitrate cluster around a different line with about three times as great a slope. Both of these lines would extrapolate to a value of E^0 about 5 mv. too high. The results for lead chloride alone approach the origin with a positive instead of the predicted negative slope. The graph of the $E^{0''}$ function clearly indicates divergence from the behavior of a typical strong electrolyte. It shows that the ionic strength alone does not determine this deviation, and suggests that the results may be treated in terms of ionic association to form one or more complexes. This was suggested sometime ago on the basis of E.M.F. data by G. Scatchard and R. H. Tefft (25), and later by E. Güntelberg (13), while additional evidence has been obtained along two very different lines. E. C. Righellato and C. W. Davies (24) have studied the anomalous conductance of lead chloride solutions, and they have interpreted the results in terms of formation of $PbCl^+$ complex ion. Independently, H. Fromherz and K. H. Lih (11) have studied the ultraviolet absorption spectra of these solutions, and also have obtained evidence for the formation of $PbCl^+$ in dilute solutions.

K FOR $PbCl^+$ FROM CONDUCTANCE

Righellato and Davies found that, over a considerable range of concentration, all strong electrolytes obeyed the equation:

$$\Lambda^0 - \Lambda\eta^{1/2} = a \cdot f(c) \tag{13}$$

Here Λ^0 and Λ are the equivalent conductances at concentrations zero and c, η is the viscosity at concentration c, a is the constant of the Debye–Hückel–Onsager equation, calculable from ionic mobilities, and $f(c)$ is the same function for all salts of the same valence type. Weak electrolytes, including lead chloride, do not obey this equation.

Righellato and Davies assumed that the conductance of the lead chloride solution is the sum of the conductances of the Pb^{++}, Cl^-, and $PbCl^+$ ions, and that the conductance of the first two ions is the same as in other (unassociated) solutions in which they are present at the same ionic strength. The conductance of the $PbCl^+$ ion they assumed to be 40 mhos, like that of all the other univalent complex ions they considered. They then determined the degree of ionic association, β, required to yield the observed conductance. The ionic concentrations in the solution are:

$$m_{Pb^{++}} = (1 - \beta)m; \quad m_{Cl^-} = (2 - \beta)m; \quad m_{PbCl^+} = \beta m$$

The actual ionic strength of the solution will be:

$$\mu' = (3 - 2\beta)m \tag{14}$$

This differs from the stoichiometric value $\mu = 3m$, which is also the actual ionic strength in a solution of a strong electrolyte. The value of β was determined by successive approximations, and the authors consider it accurate to better than 10 per cent.

Righellato and Davies then calculated the thermodynamic dissociation constant for the complex ion, given by the equation:

$$K = \frac{m_{\mathrm{Pb}^{++}} \cdot m_{\mathrm{Cl}^-}}{m_{\mathrm{PbCl}^+}} \cdot \frac{\gamma_{\mathrm{Pb}^{++}} \cdot \gamma_{\mathrm{Cl}^-}}{\gamma_{\mathrm{PbCl}^+}} \tag{15}$$

They used a form of the Debye–Hückel equation analogous to equation 11 for the individual ionic activities, namely:

$$\log \gamma_i = -Az_i^2\sqrt{\mu'} + B_i\mu' \tag{16}$$

Converting equation 15 to logarithmic form they substituted for log γ_i from 16 to obtain:

$$\log K = \log K' - A\Sigma z_i^2\sqrt{\mu'} + \Sigma B_i\mu' \tag{17}$$

TABLE 3

K for PbCl$^+$ *from conductance at 18°C.*

m	0.001	0.0025	0.005	0.01	0.025
$10^3\mu'$	2.944	6.980	13.34	24.74	53.34
$-\log K'$	1.165	1.389	1.338	1.313	1.291
$-\log K$	1.260*	1.525	1.510	1.518	1.516
K	0.0550*	0.0299	0.0309	0.0303	0.0305

Here K' is the classical (concentration) constant. In this case Σz_i^2 is 4, and A = 0.50 at 18°C.; hence rearranging gives:

$$\log K' - 2.0\sqrt{\mu'} = \log K - \Sigma B_i\mu' \tag{18}$$

A plot of the quantity on the left against μ' gives a straight line with a slope of $\Sigma B_i = 4.44$, which is larger than the corresponding value for most other salts. Comparison with equation 11 shows that $2AB\mathring{a} = \Sigma B_i$. Substituting the numerical values for the constants A and B, as tabulated by Harned and Owen (16, p. 119), we find $\mathring{a}$ = 6.66 Å. This is considerably larger than 2.75 Å., the sum of the crystal ionic radii for lead chloride (21), but somewhat more plausible than the value of 1.75 Å. required by the extended equations of LaMer, Gronwall, and Grieff. Of the fifteen a values of the alkali halides tabulated by Harned and Owen (16, p. 381) twelve exceed the sum of the crystal radii by 0.5–2 Å. and only three alkali halides fail to equal it.

Knowing ΣB_i and K' at various concentrations, Righellato and Davies calculated corresponding values of K, which are summarized in table 3. The average value of K, neglecting the uncertain starred figure at 0.001 m, is 0.0304.

K FROM ABSORPTION SPECTRA

Fromherz and Lih (11) studied the absorption spectra of lead chloride solutions from 2000 to 2750 Å. The smoothed curves of their results are shown in figure 4, where the optical density, or logarithm of the extinction coefficient, is plotted against the wave length. The extinction coefficient is defined by the Beer–Lambert law:

$$I/I_0 = 10^{-\epsilon c d} \tag{19}$$

Here I_0 and I are the initial and final intensities of light which passes through a solution of molarity c for a distance of d cm., and ϵ is the extinction coefficient, measured to 2 to 4 per cent. The curves all intersect in a single point, corresponding to an optical density of 3.740 at 2159 Å. This indicates a system of two absorbing species which they identified as Pb^{++} (hydrated) and $PbCl^+$, which probably absorbs by what E. Rabinowitch (22) calls an electron-transfer mechanism:

$$Pb^{++}Cl^- + h\nu \rightarrow Pb^+Cl \tag{20}$$

The Cl^- does not absorb appreciably in this region. Measurements on lead perchlorate solutions showed that the optical density is independent of concentration, indicating no ionic association in solutions of this salt. Since the ClO_4^- is transparent in this region, the absorption peak with a maximum optical density of 3.990 at 2085 Å. must be due to the Pb^{++}. The extinction coefficient of any dilute lead chloride solution then will be:

$$\epsilon = \beta\epsilon_{PbCl^+} + (1 - \beta)\epsilon_{Pb^{++}} \tag{21}$$

In this equation both β and ϵ_{PbCl^+} are unknowns. However, Fromherz and Lih assumed that the absorption curve for $PbCl^+$ is symmetrical, like those of $PbBr^+$ and PbI^+ which they had calculated by difference, and that it obeys the Beer–Lambert law on dilution. The curves of the two most concentrated chloride solutions were analyzed, assuming values for β_1 and β_2 over a reasonable range. Comparison of the optical density curves for $PbCl^+$ showed that the dashed curve of figure 4 was consistent with both series of measurements, giving maxima at 2270 and 2265 Å., respectively, with half-widths of 232 and 226 Å. This yielded values of $\beta_1 = 0.320$ and $\beta_2 = 0.178$. The calculated curve for $PbCl^+$ was used with the observed curves of the dilute solutions to calculate the corresponding values of β from equation 21. The results are summarized in table 4, where the maximum error in β estimated by the authors is also included.

The value of β being known, the classical ionization constant was calculated by Fromherz (10) as:

$$K' = \frac{(1 - \beta)(2 - \beta)}{\beta} c \tag{22}$$

A plot of K' against c gives the thermodynamic constant K by extrapolation and also the activity coefficient ratio. Fromherz found $K = 0.0775$ at about 22°C., with an estimated uncertainty of ±10 per cent.

Fromherz and Lih also studied solutions of lead chloride in concentrated alkali chloride solutions, and found absorption bands which they attributed to the $PbCl_4^{--}$ ion. These merge imperceptibly into the Pb^{++}–$PbCl^+$ system in dilute solution, whence they conclude no intermediate ions are formed. Rabino-

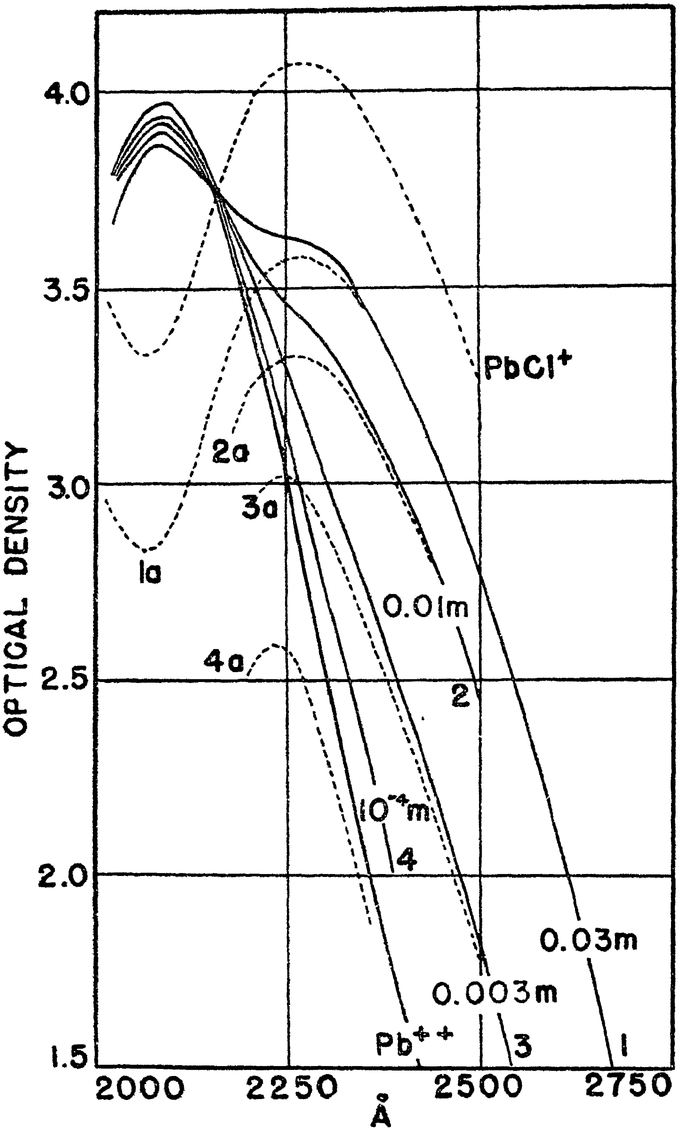


FIG. 4. The optical density of solutions of lead chloride according to Fromherz and Lih

TABLE 4

Degree of association of $PbCl^+$ *in dilute solutions of lead chloride, according to Fromherz and Lih*

10^3c	27.69	1.108	2.769	0.9229
β	0.320	0.178	0.087	0.033
$\Delta\beta$	±0.03	±0.02	±0.02	±0.008

witch (22) thinks, however, that such ions might exist, with absorption bands in the same general region as the others. In a solution 0.000167 *M* in lead chloride and 0.0167 *M* in potassium chloride they found 80 per cent Pb^{++}, 10 per cent $PbCl^+$, and 10 per cent higher complexes.

K FROM E.M.F. DATA

The large disagreement in the value of the dissociation constant of $PbCl^+$ obtained from these two methods aroused our interest in evaluating it independently from E.M.F. data. Scatchard and Tefft (25) had already stated that it was approximately 0.03, and we used their method, as modified by H. S. Harned and M. E. Fitzgerald (15). The method has been summarized by Harned and Owen (16, p. 414 *et seq.*). The E.M.F. of the cell may be expressed in terms of the degree of dissociation of the $PbCl^+$ ion, α, and the true mean activity coefficient $\gamma'_\pm$ of the lead chloride as:

$$E = E^0 - \tfrac{1}{2}k \log \alpha(1 + \alpha)^2 m^3 \gamma'^3_\pm \tag{23}$$

Following Harned and Fitzgerald, we assumed that $\gamma'_\pm$ was the same as that of the typical strong electrolyte barium chloride, as determined by E. A. Tippets and R. F. Newton (27). We transformed their Hückel-type equation for log *f* into the following:

$$\log \gamma_\pm = -\frac{1.0117\sqrt{\mu}}{1 + 1.3436\sqrt{\mu}} + 0.0388\mu \tag{24}$$

which reproduced their tabulated values of $\gamma_\pm$ with an average deviation of ±0.0004 up to 1.2 *m*. Substituting equation 24 into equation 23, we can assume values of *K* and solve for E^0 by successive approximations or, somewhat more simply, assume values of E^0 and solve for *K* by successive approximations. The appropriate equation is:

$$-\log \alpha(1 + \alpha)^2 = \left[3 \log m + 2\frac{(E - E^0)}{k}\right] \\ \frac{3.0350m^{1/2}\sqrt{2\alpha + 1}}{1 + 1.3436m^{1/2}\sqrt{2\alpha + 1}} + 0.1166m(2\alpha + 1) \tag{25}$$

From this α is determined to 0.1 per cent by two or three successive approximations. The classical dissociation constant then is calculated as:

$$K' = \alpha\frac{(1 + \alpha)}{(1 - \alpha)}m \tag{26}$$

Harned and Fitzgerald assumed that the activity coefficients of $PbCl^+$ and Cl^- are equal, and that the activity coefficient of Pb^{++} is given by the Debye–Hückel equation:

$$\log \gamma_{Pb^{++}} = -\frac{2.0234\sqrt{\mu'}}{1 + 0.3283\, \mathring{a}\sqrt{\mu'}} \tag{27}$$

where μ' is the actual ionic strength in the solution, given by equation 14, and $\mathring{a}$ is assigned the value of 5 Å., comparable to that found for barium chloride and large enough to make the extended terms of the Debye–Hückel theory negligible. Then:

$$K = K'\gamma_{\mathrm{Pb}^{++}} \tag{28}$$

We used the data of Carmody to determine K in this way, taking several values of E^0 and also varying $\mathring{a}$. The most consistent values were: $E^0 = 0.3430$ v., $\mathring{a} = 6$ Å., and $K = 0.0293$. Estimating the uncertainty in E^0 as 0.1 mv. and that in $\mathring{a}$ as 1 Å., we weighted each value of K according to the reciprocal of the sum of the resulting errors in E. The results of calculations with two different values

TABLE 5

Differences in millivolts between E.M.F. found by Carmody for the cell
Pb(*two-phase amalgam*) | ($PbCl_2$(*m*) | (AgCl(*s*)–Ag(*s*)
and that calculated by various equations

Values of constants				
E^0	0.3430	0.3429	0.3429	0.34325
$\mathring{a}$	6	7	6	1.75
K	0.0293	0.0280	0.0275	∞
m	$\Delta E = E$ (calculated) $-$ E (observed) in millivolts			
0.0002	0.05	−0.04	−0.02	−0.25
0.0006	0.02	−0.04	−0.03	−0.05
0.0010	0.03	−0.02	0.02	−0.05
0.0013	0.01	−0.02	0.00	0.05
0.0023	−0.02	0.02	0.07	0.15
0.0026	0.11	0.14	0.16	0.05
0.0052	0.00	0.09	0.14	0.05
0.0104	−0.13	0.11	0.09	0.05
0.0205	−0.43	0.00	−0.11	
0.0295	−0.61	−0.08	−0.24	
0.0391	−0.96	−0.36	−0.53	
Av. (1–8)	±0.05	±0.06	±0.07	±0.09
(1–10)	±0.14	±0.06	±0.09	
(1–11)	±0.21	±0.08	±0.13	

of E^0 and $\mathring{a}$ are shown in table 5. The tabulated values of K are based upon the weighted values of the first eight experiments, up to and including the 0.01 m solution. Above this concentration the decreasing value of K, which can be compensated somewhat by increasing $\mathring{a}$, probably is due to formation of higher complexes in addition to $PbCl^+$, rather than to a decrease in $\gamma_{\mathrm{Pb}^{++}}$. We have included also the deviations which LaMer, Gronwall, and Grieff calculated from these same data. A comparison shows that the assumption of complex ion formation is slightly more consistent with Carmody's data than the extended Debye–Hückel theory, although deviations below 0.05 mv. are not significant.

An analysis of Hannan's results at 25°C. showed that, over the range covered (0.0005–0.010 m), they were most consistent with values of $E^0 = 0.3433$, $\mathring{a} = 6$, and $K = 0.0289$. The average deviation in E was ±0.16 mv., chiefly because the value at the highest concentration was 0.37 mv. low. The result for E^0 is consistent with Hannan's statement that Carmody's $E^{0'}$ values were about 0.3 mv. higher than hers, a result which she attributed to his use of amalgams made from stick lead rather than electrolyzed amalgams.

The values of E^0 for the cell with lead chloride as electrolyte are summarized in table 6. The method of calculation based on the formation of $PbCl^+$ gives results about midway between those of direct linear extrapolation of $E^{0'}$ against $\mu^{1/2}$ and the extended theory of LaMer, Gronwall, and Grieff.

Since the picture of incomplete ionization of $PbCl^+$, with a constant of about 0.029 at 25°C., fitted the results of Carmody and Hannan for pure lead chloride solutions so satisfactorily, we tried to apply it also to the mixed solutions. If m_1, m_2, and m_3 are respectively the stoichiometrical concentrations of lead chloride, alkali chloride, and barium nitrate, the equation for the cell in terms of the degree of dissociation of the $PbCl^+$, analogous to equation 23 is:

TABLE 6

A comparison of E^0 values from various sources for the cell
Pb(*two-phase amalgam*) | $PbCl_2(m)$ | AgCl(*s*)–Ag(*s*)

METHOD OF EVALUATION		DATA	
		Carmody	Hannan
Linear *vs.* $\mu^{1/2}$	Graphical	0.3426	
	Least squares	0.3427	0.3429
K for $PbCl^+$	This article	0.3430	0.3433
	Scatchard and Tefft	0.3430	
LaMer, Gronwall, and Grieff		0.34325	0.3435

$$E = E^0 - \tfrac{1}{3}k' \log \alpha m_1[(1 + \alpha)m_1 + m_2]^2 - k' \log \gamma'_\pm \tag{29}$$

where $\gamma'_\pm$ is the true mean activity coefficient of the dissociated lead chloride. If we assume that this is the same as $\gamma_\pm$ for barium chloride at 25°C., given by equation 24, we can substitute it into equation 29 and obtain:

$$E + \tfrac{1}{3}k' \log \alpha m_1[(1 + \alpha)m_1 + m_2]^2 = E^0 + \frac{1.0117k'\sqrt{\mu'}}{1 + 1.3436\sqrt{\mu'}} - 0.0388k'\mu' \tag{30}$$

However, for purposes of correlation, it is more convenient to substitute for $\gamma'_\pm$ in equation 29 the limiting law for a 2–1 electrolyte, and define a function $E^{0'''}$ as:

$$E^{0'''} \equiv E + \tfrac{1}{3}k' \log \alpha m_1[(1 + \alpha)m_1 + m_2]^2 - 1.0117k'\sqrt{\mu'} \tag{31}$$

The value of α is obtained by successive approximations from the equation:

$$\alpha = \frac{1}{2}\left[\sqrt{\left(1 + \frac{m_2}{m_1} + \frac{K}{m_1\gamma'_{Pb^{++}}}\right)^2 + \frac{4K}{m_1\gamma'_{Pb^{++}}}} - \left(1 + \frac{m_2}{m_1} + \frac{K}{m_1\gamma'_{Pb^{++}}}\right)\right] \tag{32}$$

where $\gamma'_{\mathrm{Pb}^{++}}$ is obtained from the Debye–Hückel equation (equation 27) putting $\mathring{a} = 6$:

$$\log \gamma'_{\mathrm{Pb}^{++}} = -\frac{2.0233\sqrt{\mu'}}{1 + 1.9699\sqrt{\mu'}} \tag{33}$$

and (34)

$$\mu' = (2\alpha + 1)m_1 + m_2 + 3m_3$$

The values of $E^{0'''}/k'$ thus obtained, setting $K = 0.029$, are listed in the eighth column of table 2. When plotted against the values of μ' listed in the ninth column of the same table, they all fall close to a single line, as shown in figure 5. Comparison of equations 30 and 31 shows that, if the mean activity coefficient of the dissociated lead chloride is the same as that of barium chloride:

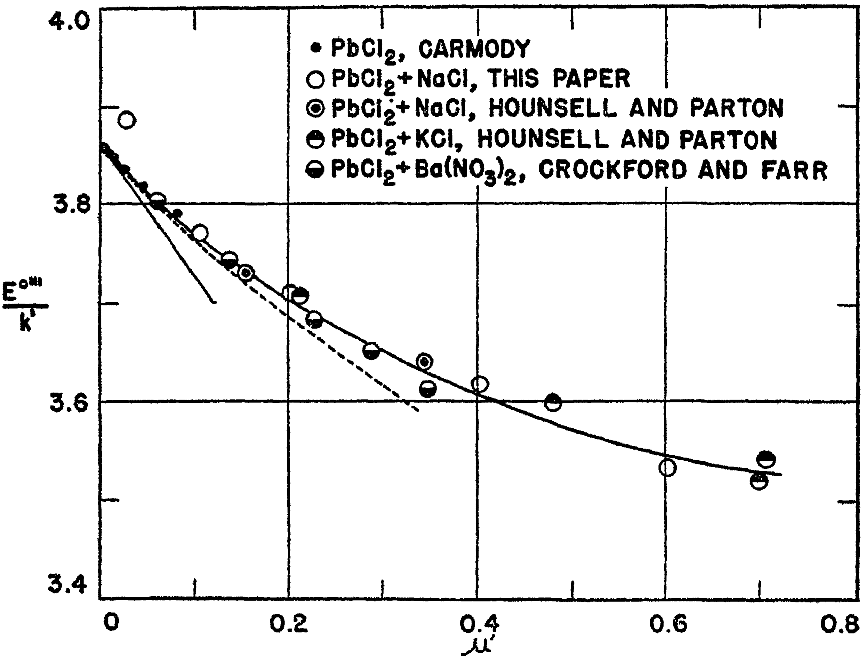


FIG. 5. $E^{0'''}/k'$ (based on $K = 0.0290$ for $PbCl^+$ and $\mathring{a} = 6$) for lead chloride alone and with added electrolytes at 25°C.

$$E^{0'''}/k' = E^0/k' - \frac{1.3593\mu'}{1 + 1.3436\sqrt{\mu'}} - 0.0388\mu' \tag{35}$$

The straight line of figure 5 indicates the limiting slope of this curve and the dashed line shows its further course. The calculated curve agrees quite well with the experimental in dilute solution, showing that the value of $\gamma'_{\pm}$ is about that of barium chloride, but the divergence is appreciable above 0.1 μ. This is not surprising, considering the assumptions involved. Thus the neglect of higher complexes in the more concentrated solutions would tend to make the values of α

from equation 32 too large and may explain some of the systematic deviations. The fact that the results for added chlorides fall slightly above those for added barium nitrate also may be explained in the same way, or may be due to incomplete dissociation of the latter salt, as suggested by Dr. Redlich. However, the differences are rather small, and this single plot serves to correlate all the experimental results. When it is used for extrapolation to obtain E_0, the value thus obtained agrees within 0.3 mv. with that of Hannan, so that it is much more suitable for extrapolation than $E^{0''}$.

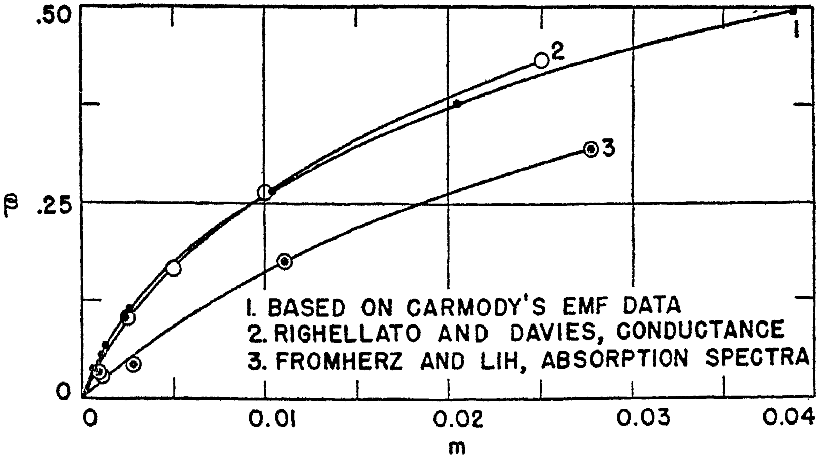


FIG. 6. Fraction (β) of lead chloride associated to form $PbCl^+$ at 25°C., calculated by three independent methods.

TABLE 7

The dissociation constant of $PbCl^+$ *calculated by various methods*

AUTHOR	METHOD	K
This paper	Carmody's E.M.F. data	0.0293
This paper	Hannan's E.M.F. data	0.0289
Righellato and Davies	Conductance	0.0304
Redlich	von Ende's solubility data	0.06
Fromherz and Lih	Absorption spectra	0.08
Güntelberg	E.M.F. data	0.10

SUMMARY OF VALUES OF K

The results of the calculations of the dissociation constant of $PbCl^+$ based on E.M.F. measurements show excellent agreement with the analysis of conductance data by Righellato and Davies, and wide divergence from the conclusions of Fromherz and Lih based on absorption spectra. This appears clearly in figure 6, where we have plotted the fraction of lead associated to form $PbCl^+$ at different concentrations. In table 7 we have listed the values of the dissociation constant of $PbCl^+$ discussed above, along with several others. C. L. von Ende (28) determined the freezing point of lead chloride solutions and calculated the concentration of Pb^{++} and $PbCl^{+}$ in the saturated solution at 25°C., assuming that the ions are ideal solutes. Neglecting activity coefficients, O. Redlich (23) estimated the dissociation constant on the basis of these data, although he does not consider that the figure should carry any weight, because of the assumptions involved. The value given by E. Güntelberg (13) is based on his measurement of the E.M.F. of cells containing lead chloride alone and in the presence of hydrochloric acid or potassium chloride. Since we have seen only the abstract of his work we do not have the details of experimentation or computation, but his results seem at variance with the careful and concordant work of Carmody and Hannan. We have no explanation for the wide discrepancy between the results of about 0.03 based on these E.M.F. data and conductance data, and those of about 0.09, based upon absorption spectra and Güntelberg's data. Further work is needed to clear up this discrepancy.

REFERENCES

(1) ALLMAND, A. J., AND BURRAGE, L. J.: Trans. Faraday Soc. **23**, 470 (1927).
(2) ALLMAND, A. J., AND BURRAGE, L. J.: Trans. Faraday Soc. **29**, 679 (1933).
(3) ALLMAND, A. J., AND HUNTER, E.: Trans. Faraday Soc. **24**, 300 (1928).
(4) BRASHER, D. M., AND PARTON, H. N.: Trans. Faraday Soc. **31**, 681 (1935).
(5) BRÖNSTED, J. N.: Z. physik. Chem. **56**, 645 (1906).
(6) CARMODY, W. R.: J. Am. Chem. Soc. **51**, 2905 (1929).
(7) CROCKFORD, H. D., AND FARR, H. O., JR.: J. Am. Chem. Soc. **58**, 87 (1936).
(8) CROCKFORD, H. D., AND THOMAS, H. C.: J. Am. Chem. Soc. **55**, 568 (1933).
(9) DOLE, M.: *The Glass Electrode*, Chap. 5. John Wiley and Sons, Inc., New York (1941).
(10) FROMHERZ, H.: Z. physik. Chem. **153A**, 376 (1931).
(11) FROMHERZ, H., AND LIH, K. H.: Z. physik. Chem. **153A**, 321 (1931).
(12) GARRELS, R. M.: Econ. Geol. **36**, 729 (1941).
(13) GÜNTELBERG, E.: Kem. Maanedsblad **19**, 85 (1938); Chem. Abstracts **34**, 6507 (1940).
(14) HANNAN, SISTER M. C.: "Some Thermodynamic Properties of Lead Chloride in Aqueous Solution," Ph.D. Dissertation, The Catholic University of America, 1936.
(15) HARNED, H. S., AND FITZGERALD, M. E.: J. Am. Chem. Soc. **58**, 2624 (1936).
(16) HARNED, H. S., AND OWEN, B. B.: *The Physical Chemistry of Electrolytic Solutions.* Reinhold Publishing Corporation, New York (1943).
(17) HOUNSELL, E. R., AND PARTON, H. N.: Trans. Faraday Soc. **33**, 629 (1937).
(18) HUNTER, E., AND ALLMAND, A. J.: Trans. Faraday Soc. **32**, 502 (1936).
(19) LAMER, V. K., GRONWALL, T. H., AND GRIEFF, L. J.: J. Phys. Chem. **35**, 2245 (1931).
(20) PARTON, H. N.: Trans. Faraday Soc. **31**, 686 (1935).
(21) PAULING, L.: *The Nature of the Chemical Bond*, p. 326. Cornell University Press, Ithaca, New York (1939).
(22) RABINOWITCH, E.: Rev. Modern Phys. **14**, 112 (1942).
(23) REDLICH, O.: Chem. Revs. **39**, 333 (1946).
(24) RIGHELLATO, E. C., AND DAVIES, C. W.: Trans. Faraday Soc. **26**, 592 (1930).
(25) SCATCHARD, G., AND TEFFT, R. F.: J. Am. Chem. Soc. **52**, 2272 (1930).
(26) SCATCHARD, G., PRENTISS, S. S., AND JONES, P. T.: J. Am. Chem. Soc. **54**, 2690 (1932).
(27) TIPPETS, E. A., AND NEWTON, R. F.: J. Am. Chem. Soc. **56**, 1675 (1934).
(28) VON ENDE, C. L.: Z. anorg. Chem. **26**, 129 (1901).

AQUEOUS SOLUTIONS OF PARAFFIN-CHAIN SALTS[1]

E. C. LINGAFELTER

Department of Chemistry and Chemical Engineering, University of Washington, Seattle, Washington

Received October 25, 1948

It is well recognized that the behavior of aqueous solutions of paraffin-chain salts is to be explained by the appearance of colloidal aggregates. However, most of the questions relating to the details of the aggregation are still unsettled.

Some of the experimental results are reviewed and their bearing on these questions is discussed.

Aqueous solutions of paraffin-chain salts form an interesting case in which the effects of dissociation and association are clearly evident. At low concentrations these salts behave as ordinary strong electrolytes. As the concentration is increased a point is reached, at a concentration dependent mainly upon the size and shape of the paraffin chain and the magnitude of the electrical charge of the ions, where the paraffin-chain ions more or less suddenly aggregate into particles of colloidal dimensions. This aggregation of the paraffin-chain ions in aqueous solution into colloidal micelles is well known (12, 23). However, the details of the phenomenon are as yet incompletely elucidated. The uncertainty as to the detailed interpretation of the properties of the solutions is due to at least three factors: (*1*) the systems are undoubtedly complex, (*2*) the magnitude of the interionic effects is unknown, and (*3*) the various data which have been reported often disagree considerably.

The questions which are to be answered in the complete description of these solutions are: (*a*) Is there at all concentrations only one type of micelle, or are there several, such as the "ionic micelle", the "lamellar micelle", etc.? (*b*) For a given type of micelle is there a fairly well-defined optimum size, or is there a particle-size distribution with a broad maximum? (*c*) Do the micelles consist solely of paraffin-chain ions, or are some of the simple ions of opposite charge (the "gegen ions" or "counter ions") included in the aggregate? (*d*) How large are the micelles, and how many ions of each type are included in one colloidal particle? (*e*) How are the constituent ions arranged in the micelle? (*f*) What are the concentrations of the free simple ions in equilibrium with the micelles? Obviously, all of these questions cannot be answered by thermodynamic considerations alone but require the application of all types of measurement which may be available.

McBain, who first suggested that the behavior of these solutions could be explained in terms of the formation of colloidal aggregates, has maintained that there are at least two, quite different, types of micelles (23). His "ionic" micelle is considered to be a spherical aggregate of not more than ten paraffin-chain ions, retaining its full electrical charge, while the "lamellar" micelle is a much larger particle, having the paraffin-chain ions in an arrangement similar to that in the solid state (3, 42), with a low electrical charge due to the inclusion of a considerable number of gegen ions. Ralston (34), Van Rysselberghe (40) and Stauff (37) have also considered that there must be two types of micelles.

---

[1] Presented at the Symposium on Thermodynamics and Molecular Structure of Solutions, which was held under the auspices of the Division of Physical and Inorganic Chemistry at the 114th Meeting of the American Chemical Society, Portland, Oregon, September 13 and 14, 1948.

McBain, Ralston, and Van Rysselberghe have based their arguments on the assumption that the interionic forces in the solutions are small, i.e., of the same order as in solutions of ordinary uni-univalent strong electrolytes. McBain has argued that, since the ionic charges on the micelle surface are separated from each other by several atomic diameters, their electrostatic effects will be essentially independent. As experimental evidence supporting this assumption, McBain and Searles (25) and McBain and Brady (24) have measured freezing points and conductivities of solutions containing both paraffin-chain salts and simple salts and have found that the effects are essentially additive. Some doubt may be cast on the interpretation of these data by the fact that the number, size, and electrical charge of the micelles may be changed by the presence of the added salt, which may lead to a set of compensating factors. With this assumption of small interionic effects, it has not been found possible to explain the observed properties of the solutions without the use of several types of micelles.

Hartley, on the other hand, believes that the properties of the solutions are explainable with the assumption of only one type of micelle (12), if one takes into account the large interionic effects due to the high electrical charge of the micelles. As experimental evidence of the existence of large interionic effects, he refers to measurements of electrical conductivity at high frequencies (36) and at high field strengths (19). Both of these investigations, while not conclusive, indicate the existence of large interionic effects. Similarly, Stone (39) has shown that sodium tetradecane-α,ω-disulfonate in aqueous solution gives an ionic strength of a uni-bivalent electrolyte. He has suggested that the bivalent behavior of these ions may be due to a bending of the paraffin chain, bringing the two ionic groups into close proximity. That this is not necessarily true is shown by unpublished data from this laboratory on the effect of sodium 4,4′-biphenyldisulfonate (15) on the solubility of thallous chloride. Although the sulfonate groups are about 11 Å., apart and the ion is rigid, the 4,4′-biphenyldisulfonate ion behaves as a bivalent ion. The recent work of Fuoss and Strauss (9) and Edelson and Fuoss (8) on polyelectrolytes also clearly shows the influence of large interionic effects.

None of the results relating to the magnitude of the interionic effects so far published are conclusive, and it is to be hoped that additional experimental data may appear to resolve the dilemma. Two lines of investigation which may yield significant results are now under way in our Laboratories. First, one can prepare polymers of known molecular weight, containing known numbers and distributions of ionic groups. Studies of the ionic strength effects of salts containing these ions will be of considerable interest. Secondly, to avoid the problem of unknown concentrations arising from the adherence of gegen ions to the micelle, one can attempt to measure the effective ionic strength of solutions of paraffin-chain salts by investigating the rate of a reaction between two ions bearing electrical charges of the same sign as that of the micelles.

Stauff has inferred the presence of two types of micelles, mainly from x-ray diffraction studies (37). At moderate to high concentrations a long-spacing diffraction band appears, which Stauff and others have attributed to the presence of lamellar micelles. As the concentration of paraffin-chain salt is decreased, the spacing increases (interpreted as inclusion of water between the layers of paraffin-chain ions) and the intensity of the band decreases. By plotting the diffraction intensity against the concentration and extrapolating to zero intensity, Stauff concludes that the lamellar micelle exists only at high concentrations and that the colloidal effects observed at lower concentrations are due to a small spherical micelle. Dervichian (6, 7) has suggested, however, that there is only one type of micelle and that the long-spacing diffraction band is due to an intermicelle spacing which only appears after the concentration of micelles reaches a certain limiting value. Mattoon, Stearns, and Harkins (20, 21) agree with Dervichian and present as evidence a newly discovered long-spacing band which they attribute to the single micelles.

The nature of the particle-size distribution curve has been considered theoretically by Meyer and van der Wyk (26), assuming a parallel arrangement of the paraffin chains and neglecting electrostatic effects. They calculate a curve having a very broad, flat maximum which shifts to larger sizes as the concentration increases. On the other hand, Hartley (12) has pointed out the greater likelihood of a spherical micelle with an essentially liquid-like arrangement of the paraffin chains in the interior. The assumption of the spherical micelle leads to a rather sharply defined optimum size, the radius of the sphere being approximately equal to the length of a single, fully extended paraffin-chain ion. Diffusion (1, 10, 14, 41) and ultracentrifuge (28) measurements indicate micellar radii which are in agreement with expectation from the spherical micelle. Thus, the direct experimental evidence as to the size of the micelles indicates, at least at concentrations just above the critical concentration, a micelle radius approximately equal to the length of the paraffin-chain ion. A micelle of this size, having an interior density equal to that of a liquid paraffin hydrocarbon, would contain about 50 ions in the case of a twelve-carbon chain and about 90 for a sixteen-carbon chain.

The sharpness of the change in properties at the critical concentration indicates (29) that the micelles must be fairly large, with a rather small fraction of aggregates of intermediate size.

Direct determination of the mean ionic activity of 1-dodecanesulfonic acid by E.M.F. measurements (30, 31, 32, 43) shows the behavior of an ordinary strong electrolyte at concentrations below the critical concentration (about 0.008 M). Above the critical concentration the mean ionic activity is practically constant, rising only very slowly with increasing concentration.

The evidence for the inclusion of gegen ions in the micelle is considerable.

Aggregation of ions all of one sign would lead to an increase in conductance (22), whereas the equivalent conductance is usually observed to drop above the critical concentration. The transference number of the gegen ion is sometimes negative (13), owing to the transference of attached gegen ions by the micelle in the opposite direction from their motion when free. The yellow color of aqueous solutions of cetylpyridinium iodide (12) can hardly be due to anything but a firm attachment between the ions. The actual fraction of gegen ions attached has not been determined accurately, but estimates range from 50 to 80 per cent.

The evidence concerning the arrangement of the paraffin-chain ions in the micelle is so far inconclusive. Hartley (12) has argued for a random arrangement similar to liquid paraffin both on the basis of probability and from consideration of the lack of specificity in the solubilization of water-insoluble organic molecules by aqueous solutions of paraffin-chain salts. Similarly, Corrin (4) has found that the relationship between the amount of oil dissolved and the increase in x-ray spacing is in somewhat better agreement with a spherical micelle than the cylindrical model. The small-angle x-ray diffraction patterns obtained from solutions containing micelles have been interpreted as evidence for a lamellar micelle with parallel arrangement of the paraffin chains (21, 23, 37), and Brady (2) has discussed the interpretation of the shift of position of the long-spacing band with concentration in terms of the application of the theory of the diffuse ionic double layer to the lamellar micelle. However, Corrin (5) has recently shown that it is possible to account for the observed diffraction effects with the assumption of spherical micelles.

The concentrations of the free ions in equilibrium with the micelles cannot be accurately determined until we are able to take the interionic effects into account by appropriate activity coefficients. The estimates which have been made do not in all cases agree. For example, Hartley's (11) data on the solubilization of *trans*-azobenzene in aqueous solutions of cetylpyridinium sulfate indicate that the concentration of free cetylpyridinium ions decreases above the critical concentration. The data of Stearns, Oppenheimer, Simons, and Harkins (38) on the solubilization of 2-nitrodiphenylamine in aqueous solutions of potassium laurate can be interpreted in the same manner. On the other hand, the solubilization data for dimethylaminoazobenzene in aqueous solutions of potassium laurate (17) have been interpreted by Kolthoff and Stricks to indicate that the concentration of laurate ions does not decrease above the critical concentration. Similarly, the electromotive force data of Kolthoff and Johnson (16) seem to indicate that the activity of laurate ions in potassium laurate solutions is constant above the critical concentration, while the electromotive force measurements of Walton on solutions containing hydrochloric acid and 1-dodecanesulfonic acid (44) or hydrochloric acid and dodecylbenzyldimethylammonium chloride (45) indicate that the activity of the paraffin-chain ion decreases above the critical concentration. In this connection it should be pointed out that any equilibrium constant calculations involving a micelle which contains more paraffin-chain ions than gegen ions leads to a decrease of activity of the paraffin-chain ion above the critical concentration.

One further thermodynamic problem has arisen in the application of the Gibbs adsorption theorem to the surface tensions of aqueous solutions of paraffin-chain salts. Most of the reported data have shown the expected rapid decrease in surface tension with increasing concentration. At the critical concentration, however, the surface tension passes through a minimum and then rises considerably with further increase in concentration (18, 33). This rise would seem to indicate either a negative adsorption or a decrease in activity, both of which are extremely unlikely. Many hypotheses have been proposed to reconcile the anomalies (*cf.* 23), but recently Miles and Shedlovsky (27) and Reichenberg (35) have shown that at least in some cases, and probably in all, the minimum has been caused by traces of impurities (long-chain alcohols, paraffin-chain salts with polyvalent gegen ions, etc.) which are more highly adsorbed than the substance investigated, but which are dissolved in the micelles above the critical concentration. When these impurities are removed, the surface tension becomes practically constant above the critical concentration, as would be expected from the activity data.

REFERENCES

(1) Abbott, A. D.: Ph.D. Thesis, University of Washington, 1942.

(2) Brady, A. P.: Paper No. 4 presented before the Division of Colloid Chemistry at the 114th Meeting of the American Chemical Society, Portland, Oregon, September, 1948.

(3) Buerger, M. J.: Am. Mineral. **30,** 551 (1945).

(4) Corrin, M. L., quoted by Mattoon, R. W., Stearns, R. S., and Harkins, W. D.: J. Chem. Phys. **16,** 644 (1948).

(5) Corrin, M. L.: J. Chem. Phys. **16,** 844 (1948).

(6) Dervichian, D.: J. chim. phys. **38,** 59 (1941).

(7) Dervichian, D., and Lachampt, F.: Bull. soc. chim. [5] **12,** 189 (1945).

(8) Edelson, D., and Fuoss, R. M.: J. Am. Chem. Soc. **70,** 2832 (1948).

(9) Fuoss, R. M., and Strauss, U. P.: J. Polymer Sci. **3,** 246 (1948).

(10) Hakala, N. V.: Ph.D. Thesis, University of Wisconsin, 1943.

(11) Hartley, G. S.: J. Chem. Soc. **1938,** 1968.

(12) Hartley, G. S.: Kolloid-Z. **88,** 22 (1939).

(13) Hartley, G. S., Collie, B., and Samis, C. S.: Trans. Faraday Soc. **32,** 795 (1936).

(14) Hartley, G. S., and Runnicles, D. F.: Proc. Roy. Soc. (London) **A168,** 420 (1938).

(15) Henniker, C. J. C.: Ph.D. Thesis, University of Washington, 1943.

(16) Kolthoff, I. M., and Johnson, W.: J. Phys. Colloid Chem. **52,** 22 (1948).

(17) Kolthoff, I. M., and Stricks, W.: J. Phys. Colloid Chem. **52,** 915 (1948).

(18) Lingafelter, E. C., Wheeler, O. L., and Tartar, H. V.: J. Am. Chem. Soc. **68,** 1490 (1946).

(19) Malsch, J., and Hartley, G. S.: Z. physik. Chem. **A170,** 321 (1934).

(20) Mattoon, R. W., Stearns, R. S., and Harkins, W. D.: J. Chem. Phys. **15,** 209 (1947).

(21) Mattoon, R. W., Stearns, R. S., and Harkins, W. D.: J. Chem. Phys. **16,** 644 (1948).

(22) McBain, J. W.: Trans. Faraday Soc. **9,** 99 (1913).

(23) McBain, J. W.: In *Colloid Chemistry,* J. Alexander (*Editor*), Vol. 5, p. 102. D. Van Nostrand Company, Inc., New York (1944).

(24) McBain, J. W., and Brady, A. P.: J. Am. Chem. Soc. **65,** 2072 (1943).

(25) McBain, J. W., and Searles, J.: J. Phys. Chem. **40,** 493 (1936).

(26) Meyer, K. H., and van der Wyk, A.: Helv. Chim. Acta **20,** 1321 (1937).

(27) Miles, G. D., and Shedlovsky, L.: J. Phys. Chem. **48,** 57 (1944).

(28) Miller, G. L., and Andersson, K. J. I.: J. Biol. Chem. **144,** 475 (1942).

(29) Murray, R. C., and Hartley, G. S.: Trans. Faraday Soc. **31**, 183 (1935).
(30) Neff, L. L.: Ph.D. Thesis, University of Washington, 1943.
(31) Neff, L. L., Tartar, H. V., and Lingafelter, E. C.: Paper No. 3 presented before the Division of Colloid Chemistry at the 108th Meeting of the American Chemical Society, New York City, September, 1944.
(32) Neff, L. L., Wheeler, O. L., Tartar, H. V., and Lingafelter, E. C.: J. Am. Chem. Soc. **70**, 1989 (1948).
(33) Powney, J., and Addison, C. C.: Trans. Faraday Soc. **33**, 1243 (1937).
(34) Ralston, A. W.: Ann. N. Y. Acad. Sci. **46**, 351 (1946).
(35) Reichenberg, D.: Trans. Faraday Soc. **43**, 467 (1947).
(36) Schmid, G., and Larsen, E. C.: Z. Elektrochem. **44**, 651 (1938).
(37) Stauff, J.: Kolloid-Z. **89**, 224 (1939).
(38) Stearns, R. S., Oppenheimer, H., Simon, E., and Harkins, W. D.: J. Chem. Phys. **15**, 496 (1947).
(39) Stone, G. C. H.: J. Am. Chem. Soc. **62**, 572 (1940).
(40) Van Rysselberghe, P.: J. Phys. Chem. **48**, 62 (1944).
(41) Vetter, R. J.: J. Phys. Colloid Chem. **51**, 262 (1947).
(42) Vold, R. D., and Vold, M. J.: In *Colloid Chemistry*, J. Alexander (*Editor*), Vol. **5**, p. 266. D. Van Nostrand Company, Inc. New York (1944).
(43) Walton, H. F.: J. Am. Chem. Soc. **68**, 1180 (1946).
(44) Walton, H. F.: J. Am. Chem. Soc. **68**, 1182 (1946).
(45) Walton, H. F., Hiebert, E. N., and Sholtes, E. H.: J. Colloid Sci. **1**, 385 (1946).

THE COMPUTATION OF THE THERMODYNAMIC PROPERTIES OF REAL GASES AND MIXTURES OF REAL GASES[1]

JAMES A. BEATTIE

Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts

Received October 25, 1948

CONTENTS

I. Introduction 142
II. Definition and symbols 143
III. The method of van der Waals 144
IV. The Gibbs–Dalton law 145
V. Thermodynamic relations for gas mixtures from the Gibbs–Dalton law 148
VI. Application of the Gibbs–Dalton law to mixtures of real gases 150
VII. The Lewis and Randall rule 151
VIII. Thermodynamic relations for gas mixtures from the Lewis and Randall rule 153
IX. Application of the Lewis and Randall rule to mixtures of real gases 155
  A. Compressibility of gas mixtures 155
  B. Effect of an inert gas on univariant equilibria 156
  C. Chemical equilibria 156
  D. Liquid–vapor equilibria 158
X. Comparison of the Gibbs–Dalton law and the Lewis and Randall rule 158
XI. The general limit method 159
XII. The general limit method for pure gases 161
  A. The assumptions 161
  B. The integration path 162
  C. Derivations for the independent variables p and T 163
  D. Derivations for the independent variables V and T 164
  E. Equations for pure gases in terms of p and T 165
  F. Equations for pure gases in terms of V and T 166
XIII. The general limit method for gas mixtures 166
  A. The assumption 166
  B. Relations for the region of low pressure 167
  C. Relations for the region of very low pressure 168
  D. General relations 169
  E. Equations for gas mixtures in terms of p and T 170
  F. Equations for gas mixtures in terms of V and T 171
  G. General equilibrium relationships 171
XIV. Integration of the equations 173
  A. The equation of state for pure gases 173
  B. Relations among the equation of state constants and evaluation of the constants from critical data 176
  C. The equation of state for gas mixtures 178
  D. Integrated approximate equations for pure gases 180
  E. Integrated approximate equations for gas mixtures 181

---

[1] Presented at the Symposium on Thermodynamics and Molecular Structure of Solutions, which was held under the auspices of the Division of Physical and Inorganic Chemistry at the 114th Meeting of the American Chemical Society, Portland, Oregon, September 13 and 14, 1948.

F. Integrated approximate equilibrium relationships 183
G. The energy and entropy constants 186
XV. Comparison of the general limit method with the Lewis and Randall rule and the Gibbs–Dalton law 187
XVI. References 188

I. INTRODUCTION

It is well known that the thermodynamic properties of a pure gas can be computed with reference to a particular standard state from the usual thermodynamic differential equations, and the result expressed in terms of elementary functions or presented in an equivalent tabular or graphical form provided (*1*) the heat capacity at constant pressure [or volume] is known as a function of temperature for some one pressure [or density], and (*2*) there is given an equation of state with determined values of the parameters, or there are available extensive pressure–volume–temperature or equivalent data over the entire range of temperatures and pressures involved.

Moreover, the properties of a gas mixture of invariant composition can be evaluated with reference to a particular standard state of the mixture from the same relations used for pure gases provided thermal and compressibility data on that particular mixture are given. Such a computation is in general pointless.

The problem of importance to be solved in the treatment of the thermodynamic properties of gas mixtures, including mixtures in which chemical reactions may occur, is the computation of the properties of all possible mixtures (with reference to a particular standard state of the constituent gases) from thermal and compressibility data on the pure constituents alone, and a knowledge of certain integration constants of the pure gases. The solution of the problem involves the use of both thermodynamic and statistical methods. Each involves certain assumptions. Since those underlying the thermodynamic discussion can be stated more explicitly and have a more direct physical basis than those on which the statistical treatment rests, it is important to separate the two methods until near the end of the argument when the definite integrals resulting from the thermodynamic discussion are evaluated by use of relations obtained from statistical considerations.

Five methods of attack have been proposed:

1. The ideal gas formulation.
2. The method of van der Waals (92, 100, 181).
3. The Gibbs–Dalton law (61).
4. The Lewis and Randall rule (125).
5. The general limit method (73).

Gillespie (66) suggested that the term "equilibrium pressure" of a gas in a mixture be used to denote the pressure of a pure gas in isothermal equilibrium with a gas mixture through a membrane permeable to that gas alone, and that "partial pressure" of a gas in a mixture be reserved for the product of the mole fraction of the gas in the mixture into the total pressure of the mixture.

In terms of these definitions the ideal gas formulation assumes that (*1*) the equilibrium pressure of a gas in a mixture is equal to its partial pressure, and (*2*) the equation of state of a gas mixture is $pV = \Sigma_i n_i RT$. Instead of (*2*) we may simply assume that the energies of the constituent gases are functions of temperature only (15). The relations for the thermodynamic properties of gas mixtures resulting from this formulation are the familiar ones found in textbooks on thermodynamics, and are the ones approached by the equations given by any of the four other methods as the pressure on the mixture approaches zero or as the equation of state constants are allowed to approach zero.

The four other methods will be considered in detail later.

II. DEFINITIONS AND SYMBOLS

Let p, V, T be the total pressure, total volume, and Kelvin temperature and E, S, H, F_{VT}, F_{pT}, and C_V, and C_p be the total energy, entropy, heat content, T-V thermodynamic potential (work content), T-p thermodynamic potential (free energy), and constant-volume and constant-pressure heat capacities, respectively, of n moles of a pure gas or of $\Sigma_i n_i$ moles of a gas mixture consisting of $n_1, n_2, \cdots$ moles of the gases $1, 2, \cdots$ Let μ be the chemical potential of a pure gas and μ_i the chemical potential of gas i in a mixture.

Let $C = \Sigma_i n_i/V$ be the concentration of the mixture; and let $C_i = n_i/V$, $x_i = n_i/\Sigma_i n_i$, and f_i be respectively the concentration, mole fraction, and fugacity of gas i in the mixture.

Lower case letters used for an extensive quantity denote the values for 1 mole.

If it is necessary to distinguish between a pure gas and the same gas in a mixture, the subscript k will denote the value of a quantity for the pure gas and i that for the gas in the mixture.

The subscript e denotes the value of a thermodynamic quantity for a pure gas having the same total volume, temperature, and chemical potential as in a gas mixture. Under these conditions the pure gas would evidently be in thermodynamic equilibrium with the mixture through a membrane permeable to that gas alone; and, since the volume is specified, the values of all extensive as well as intensive properties of the gas would be fixed. Such quantities will be called "equilibrium" quantities, e.g., equilibrium pressure, p_e.

The subscript c denotes the value of a thermodynamic quantity for a pure gas having the same total volume, temperature, and number of moles as in a mixture. The concentration (moles per liter) of the pure gas is the same as its concentration in the mixture. Thus p_c is the pressure exerted by a pure gas that has been separated from a mixture and has the total volume and temperature of the mixture.

The subscript P denotes the value of a thermodynamic quantity for a pure gas having the same total pressure, temperature, and number of moles as in a mixture. Thus V_P is the volume of a pure gas that has been separated from a mixture and is at the total pressure and temperature of the mixture.

Thermodynamic relations among certain of the variables are:

$$E = TS - pV + \Sigma_i \mu_i n_i$$

$$H = E + pV = TS + \Sigma_i \mu_i n_i$$

$$F_{VT} = E - TS = -pV + \Sigma_i \mu_i n_i$$

$$F_{pT} = H - TS = \Sigma_i \mu_i n_i$$

$$C_V = (\partial E/\partial T)_{Vn}$$

$$C_p = (\partial H/\partial T)_{pn}$$

where the subscript n denotes constant composition as well as constant total number of moles if for a gas mixture.

Let ν_i be the stoichiometric coefficient of substance i in the equation representing a chemical reaction, being a negative number for a reactant and positive for a product. The exact mass action law is:

$$\Sigma_i \nu_i \mu_i = 0$$

The mass action constants in terms of partial pressures, K_p, and concentrations, K_C, are defined through the relations:

$$\ln K_p = \Sigma_i \nu_i \ln px_i$$

$$\ln K_C = \Sigma_i \nu_i \ln \frac{n_i}{V}$$

In the discussion of equilibrium between condensed and gaseous solutions, we shall let y_i be the mole fraction of the i^{th} constituent of the condensed phase, and p_{0i} the vapor pressure of the pure substance at the temperature T.

III. THE METHOD OF VAN DER WAALS

In this method the particular equation of state employed is introduced early in the argument, especially in the calculation of the entropy of a gas mixture.

van der Waals (181) derived a relation for the isothermal variation of the thermodynamic potential F_{VT} for a gas mixture of constant composition by integration of $dF_{VT} = -p dV$ with his equation of state for gas mixtures:

$$p = \frac{RT}{v - b_x} - \frac{A_x}{v^2}$$

where b_x and A_x are functions of the mole fractions and the corresponding equation of state constants of the constituent gases and of certain interaction constants. During the integration the temperature and composition were held constant, so that the integration constant is evidently a function of these variables. The temperature variation was not discussed, although Kuenen (99) determined it, but the variation with composition was evaluated by passing to infinite volume and applying the equations of Gibbs (62) for the energy and entropy of a mixture of ideal gases. Although the method is open to question because of the occurrence of infinite quantities, the resulting expression is in agreement, except for the omission of certain temperature functions, with that derived from the general limit method for a van der Waals gas.

van Laar (100, 101, 102, 103, 104, 105) and van Laar and Lorenz (106, 107, 133) used the general limit method for computing the energy of a gas mixture obeying van der Waals' equation, but because of the difficulty encountered by the occurrence of an improper integral in a similar argument for the entropy they used the equation of Gibbs for the entropy of a mixture of ideal gases, replacing the volume V by $V - b_x$ to correct for the volume effect of the molecules. The term $V - b_x$ is eliminated through the van der Waals equation, so that in the final expressions only the A_x constant appears. From the energy and entropy the other thermodynamic functions of a gas mixture can be derived immediately. Although b_x and A_x are considered independent of pressure and temperature during the integration, van Laar recognized that the A constants for the constituent gases are functions of these variables and in the application of the resulting equations to data he used values appropriate to the experimental range. Application to Bartlett's (3) measurements on the system $H_2O(l)$, $H_2O(g) + N_2(g)$ gave a fair representation for the composition of the gas phase at 50°C. and 1000 atm. total pressure.

Keyes (92) applied the general limit method for the energy of a gas mixture and corrected the Gibbs expression for the entropy for the volume of the molecules, using in both cases his equation of state for gas mixtures:

$$\pi\Omega = \Sigma_i n_i RT$$

$$\pi = p + \frac{(\Sigma_i n_i A_i^{1/2})^2}{(V - \Sigma_i n_i l_i)^2}$$

$$\Omega = V - (\Sigma_i n_i \beta_i) e^{-(\Sigma_i n_i \alpha_i)/V}$$

where β_i, α_i, A_i, and l_i are equation of state constants for the pure gas i. The resulting expression for the variation of the mass action "constant" of the ammonia synthesis equilibrium with pressure gave a satisfactory agreement to 100 atm. and a tolerable agreement to 600 atm. with the empirical smoothing of the measurements of Larson and Dodge (110) and of Larson (108) by Gillespie (64, 65).

IV. THE GIBBS-DALTON LAW

As a starting point in the discussion of the thermodynamic properties of gas mixtures, Gibbs (61) proposed the following law: The pressure of a mixture of different gases is equal to the sum of the pressures of the different gases as existing each by itself at the same temperature and with the same value of its (chemical) potential.

Gillespie (72) called this the Gibbs-Dalton law and pointed out that it contained the essential idea of equilibrium necessary for a general thermodynamic treatment of gas mixtures. It is evident that the different gases "existing each by itself at the same temperature and with the same value of its potential" would each be in equilibrium with the gas mixture through a membrane permeable to that gas alone. The pressure of each pure gas under these conditions is its equilibrium pressure p_e; and the law states that, subject to the conditions:

$$\left.\begin{aligned} T \text{ (each pure gas)} &= T \text{ (mixture)} \\ V \text{ (each pure gas)} &= V \text{ (mixture)} \\ \mu_k &= \mu_i \qquad (k \equiv i = 1, 2, \cdots) \end{aligned}\right\} \quad \text{(A)}$$

the pressure p of a gas mixture is:

$$p = \Sigma_k p_{ek} \tag{1}$$

The second condition is unnecessary for equation 1 but is required to fix the value of any extensive property of the pure gases.

The results derivable from the Gibbs–Dalton law may be summarized as follows (72):

1. The concentration of a gas is the same at equilibrium on each side of a membrane permeable to it alone; and under conditions A the number of moles of each pure gas is the same as in the mixture:

$$C_i = C_{ek} \qquad (k \equiv i = 1, 2, \cdots) \tag{2}$$

$$n_i = n_{ek} \qquad (k \equiv i = 1, 2, \cdots) \tag{3}$$

2. When conditions A apply the following quantities are additive:

$$S = \Sigma_k S_{ek} \tag{4}$$

$$\left.\begin{aligned} E &= \Sigma_k E_{ek} & H &= \Sigma_k H_{ek} & F_{VT} &= \Sigma_k F_{VT_{ek}} \\ F_{pT} &= \Sigma_k F_{pT_{ek}} & C_V &= \Sigma_k C_{V_{ek}} & C_p &= \Sigma_k C_{p_{ek}} \end{aligned}\right\} \tag{5}$$

3. When conditions A apply the following are also true:

$$\left.\begin{aligned} T \text{ (each pure gas)} &= T \text{ (mixture)} \\ V \text{ (each pure gas)} &= V \text{ (mixture)} \\ n_k = n_i, \text{ that is, } C_k &= C_i \quad (k \equiv i = 1, 2, \cdots) \end{aligned}\right\} \quad \text{(B)}$$

and

$$p = \Sigma_k p_{ck} \tag{6}$$

$$\mu_i = \mu_{ck} \tag{7}$$

$$\left.\begin{aligned} & & S &= \Sigma_k S_{ck} & & \\ E &= \Sigma_k E_{ck} & H &= \Sigma_k H_{ck} & F_{VT} &= \Sigma_k F_{VT_{ck}} \\ F_{pT} &= \Sigma_k F_{pT_{ck}} & C_V &= \Sigma_k C_{V_{ck}} & C_p &= \Sigma_k C_{p_{ck}} \end{aligned}\right\} \tag{8}$$

where the subscript c denotes that conditions B apply. Equation 6 is the usual statement of Dalton's law (47, 48). Although Dalton's law is a consequence of the Gibbs–Dalton law, the converse cannot be proved, since the idea of equilibrium is lacking in Dalton's law.

4. Subject to conditions B:

$$\left(\frac{\partial p}{\partial n_i}\right)_{VT_n} = \left(\frac{\partial p_{ck}}{\partial n_k}\right)_{VT} \qquad (i \equiv k = 1, 2, \cdots) \qquad (9)$$

5. The Gibbs–Dalton law leads to a complete solution of the problem of the determination of the thermodynamic properties of gas mixtures from information on the behavior of pure gases only.

Consider a number of systems, one consisting of a gas mixture and each of the others consisting of one of the pure constituent gases. Let the conditions A apply for all variations. For the gas mixture:

$$\mathrm{d}p = (S/V)\mathrm{d}T + \Sigma_i(n_i/V)\mathrm{d}\mu_i \qquad (10)$$

and for each pure gas:

$$\mathrm{d}p_{ek} = (S_{ek}/V)\mathrm{d}T + (n_{ek}/V)\mathrm{d}\mu_k \quad (k = 1, 2, \cdots) \quad (11)$$

Thus:

$$\left(\frac{\partial p}{\partial \mu_i}\right)_{T\mu} = \frac{n_i}{V} \equiv C_i \qquad \left(\frac{\partial p}{\partial T}\right)_{\mu} = \frac{S}{V} \qquad (i = 1, 2, \cdots) \qquad (12)$$

$$\left(\frac{\partial p_{ek}}{\partial \mu_k}\right)_{T} = \frac{n_{ek}}{V} \equiv C_{ek} \qquad \left(\frac{\partial p_{ek}}{\partial T}\right)_{\mu_k} = \frac{S_{ek}}{V} \qquad (k = 1, 2, \cdots) \qquad (13)$$

The subscript μ in equation 12 denotes that all of the μ's are constant unless the differentiation is with respect to one of the μ's, in which case all of the other μ's are constant.

Let us *assume* that the total pressure of the mixture is given by the Gibbs–Dalton law. Then from equation 1:

$$\left(\frac{\partial p}{\partial \mu_i}\right)_{T\mu} = \left(\frac{\partial p_{ek}}{\partial \mu_k}\right)_{T} \qquad (i \equiv k = 1, 2, \cdots) \qquad (14)$$

$$\left(\frac{\partial p}{\partial T}\right)_{\mu} = \Sigma_k \left(\frac{\partial p_{ek}}{\partial T}\right)_{\mu_k} \qquad (15)$$

Substitution from equations 12 and 13 into equations 14 and 15 gives:

$$\frac{n_i}{V} = \frac{n_{ek}}{V}$$

$$\frac{S}{V} = \Sigma_k \frac{S_{ek}}{V}$$

Since $n/V = C$, equation 2 is proved; and since V has the same value for the gas mixture and each pure gas, equations 3 and 4 are true. To prove equation 5 we substitute from equations 1, 3, and 4 into the general thermodynamic relations for E, H, etc. given in Section 2 (page 146) and use conditions A.

From equations 2 and 3 and conditions A it is evident that conditions B hold provided the Gibbs–Dalton law is true. Now conditions B and A each define the same state of each pure gas as regards both intensive and extensive thermodynamic properties. Hence when conditions B apply the chemical potential of each pure gas has the same value as in the mixture (equation 7), and equations 1, 4, and 5 become equations 6 and 8. It should be noted that equations 7 and 8 cannot be derived from equation 6 and conditions B alone.

Differentiation of equation 6 with respect to n_i $(= n_k)$, $(i \equiv k = 1, 2, \cdots)$, subject to conditions B gives equation 9.

It is evident that equations 6, 7, and 8 give the properties of a gas mixture in terms of those of the pure constituent gases without requiring any measurements on the mixture.

V. THERMODYNAMIC RELATIONS FOR GAS MIXTURES FROM THE GIBBS–DALTON LAW

Let

$$p_k = \Phi_k(T, n_k/V) \tag{16}$$

be the equation of state for a pure gas. The Gibbs–Dalton law gives for the total pressure p of a gas mixture and for the equilibrium pressure p_{ek} of gas i in the mixture the relations:

$$p = \Sigma_k p_k = \Sigma_k \Phi_k(T, n_k/V) \tag{17}$$

$$p_{ek} = \Phi_k(T, n_k/V) \qquad (k = 1, 2, \cdots) \tag{18}$$

where conditions B apply and the summation of equation 17 extends over all of the gases in the mixture. In this relation as in those immediately following it is not necessary that the form of the function Φ be the same for all of the gases. Thus different equations of state may be used for the different pure gases, and indeed Φ can be evaluated graphically for some or all of the gases involved.

In terms of the expressions for the thermodynamic properties of pure gases given by the general limit method (see Section XII), the Gibbs–Dalton law leads to the conditions B and the following equations for gas mixtures:

$$E = \Sigma_k \left\{ \int_V^\infty \left[p_k - T \left(\frac{\partial p_k}{\partial T} \right)_{V_k} \right] \mathrm{d}V_k + n_k e_k^0 \right\} \tag{19}$$

$$H = \Sigma_k \left\{ \int_V^\infty \left[p_k - T \left(\frac{\partial p_k}{\partial T} \right)_{V_k} \right] \mathrm{d}V_k + p_k V + n_k e_k^0 \right\} \tag{20}$$

$$S = \Sigma_k \left\{ \int_V^\infty \left[\frac{n_k R}{V_k} - \left(\frac{\partial p_k}{\partial T} \right)_{V_k} \right] \mathrm{d}V_k + n_k R \ln \frac{V}{n_k R T} + n_k s_k^0 \right\} \tag{21}$$

$$F_{VT} = \Sigma_k \left\{ \int_V^\infty \left[p_k - \frac{n_k R T}{V_k} \right] \mathrm{d}V_k - n_k R T \ln \frac{V}{n_k R T} + n_k (e_k^0 - T s_k^0) \right\} \tag{22}$$

$$F_{pT} = \Sigma_k \left\{ \int_V^\infty \left[p_k - \frac{n_k R T}{V_k} \right] \mathrm{d}V_k - n_k R T \ln \frac{V}{n_k R T} + p_k V + n_k (e_k^0 - T s_k^0) \right\} \tag{23}$$

$$C_V = \Sigma_k \left\{ n_k c_{vk}^* - T \int_V^\infty \left(\frac{\partial^2 p_k}{\partial T^2} \right)_{V_k} \mathrm{d}V_k \right\} \tag{24}$$

$$C_p = \Sigma_k \left\{ n_k c_{pk}^* - n_k R - T \frac{(\partial p_k/\partial T)_{V_k}^2}{(\partial p_k/\partial V_k)_T} - T \int_V^\infty \left(\frac{\partial^2 p_k}{\partial T^2} \right)_{V_k} \mathrm{d}V_k \right\} \tag{25}$$

$$\mu_i = \int_V^\infty \left[\left(\frac{\partial p_k}{\partial n_k} \right)_{V_k T} - \frac{RT}{V_k} \right] \mathrm{d}V_k - RT \ln \frac{V}{n_k RT} + RT + (e_k^0 - Ts_k^0) \tag{26}$$

$$RT \ln f_i = \int_V^\infty \left[\left(\frac{\partial p_k}{\partial n_k} \right)_{V_k T} - \frac{RT}{V_k} \right] \mathrm{d}V_k - RT \ln \frac{V}{n_k RT} \tag{27}$$

In these relations the summations extend over all of the gases in the mixture, and:

$$e_k^0 = \int_{T_0}^T c_{vk}^* \, \mathrm{d}T + e_{0k}^0 \tag{28}$$

$$s_k^0 = \int_{T_0}^T \frac{c_{vk}^*}{T} \, \mathrm{d}T + R \ln \frac{T}{T_0} + S_{0k}^0 \tag{29}$$

The quantities e_{0k}^0 and s_{0k}^0 are the molal energy and entropy constants for the pure gas k at the temperature T_0, and c_{vk}^* the molal constant-volume heat capacity at zero pressure.

In each of the equations 19 through 25 the left-hand side of the equation is the value of a thermodynamic quantity for a gas mixture at V, T, and the mole numbers n_i ($i \equiv k = 1, 2, \cdots$). The terms enclosed in the braces on the right-hand side are to be computed for each pure gas k from equations 16 (or by equivalent graphical methods), 28, and 29 with V, T, and n_k or V, T, n_k/V, having the same values for each pure gas as in the gas mixture, and the sum taken for all of the gases composing the mixture. In equations 26 and 27 the left-hand side is the value of a quantity for gas i in the mixture, and the right-hand side is computed for the pure gas as just described.

The Gibbs–Dalton law leads directly to an expression for the mass action law in terms of concentrations. Let K_C^* be the value of K_C at zero pressure. Then from equation 26 and the condition $\Sigma_i \nu_i \mu_i = 0$ for equilibrium in a system in which chemical reaction occurs, we find:

$$\ln \frac{K_C}{K_C^*} = -\frac{1}{RT} \Sigma_k \left\{ \nu_k \left[\int_V^\infty \left[\left(\frac{\partial p_k}{\partial n_k} \right)_{V_k T} - \frac{RT}{V_k} \right] \mathrm{d}V_k \right] \right\} \tag{30}$$

$$\ln K_C^* = -\frac{1}{RT} \Sigma_k \{ \nu_k [e_k^0 + RT - Ts_k^0 + RT \ln RT] \} \tag{31}$$

To obtain $\ln K_p$ we add this term to each side of equation 30 and rearrange:

$$\ln \frac{K_p}{K_p^*} = -\frac{1}{RT} \Sigma_k \left\{ \nu_k \left[\int_V^\infty \left[\left(\frac{\partial p_k}{\partial n_k} \right)_{V_k T} - \frac{RT}{V_k} \right] \mathrm{d}V_k \right] \right\} + (\Sigma_i \nu_i) \ln \frac{pV}{(\Sigma_i n_i) RT} \tag{32}$$

$$\ln K_p^* = -\frac{1}{RT}\,\Sigma_k\{\nu_k[e_k^0 + RT - Ts_k^0]\} \qquad (33)$$

In these equations the terms inclosed in braces are evaluated for each pure gas k from equations 16, 28, and 29 with V, T, n_k/V having the same values for each pure gas as in the mixture. The summations extend over all gases in the mixture, with $\nu_k = 0$ for gases that do not take part in the chemical reaction. The last term of equation 32 is to be computed for the gas mixture, the sum $\Sigma_i n_i$ being taken for all gases in the mixture. It will be noticed that both K_C and K_p are functions of the composition of the equilibrium mixture.

To compute K_C for a given chemical reaction at a given molal volume V and temperature T: (*1*) calculate K_C^* for the given temperature and as a first approximation set $K_C = K_C^*$ (the ideal gas relation), from which a provisional set of equilibrium mole fractions can be obtained; (*2*) these mole fractions are the n_k's which together with V and T permit an approximate evaluation of the right-hand side of equation 30 and hence a second approximation to K_C and the equilibrium mole fractions. The process can be repeated until no further change in composition results. The convergence of the calculation to the final equilibrium composition is rapid.

To compute K_p for a given pressure p and temperature T from equations 32 and 33 the same method is employed. However, the volume of the mixture must be calculated at each stage of the computation from equation 17.

VI. APPLICATION OF THE GIBBS–DALTON LAW TO MIXTURES OF REAL GASES

The Gibbs–Dalton law has never been used for the calculation of the effect of pressure on chemical equilibria. Lurie and Gillespie (134) and Gillespie (72) have pointed out that the law is particularly suited to the computation of the effect of the addition of an inert gas, such as nitrogen, to a univeriant system, as NH_3(l), NH_3(g) or $BaCl_2$(s), $BaCl_2 \cdot 8NH_3$(s), NH_3(g). The calculations proceed as follows:

1. Given the total pressure p, temperature T, and composition x_1, x_2, $\cdots$ of the gas phase, to find the equilibrium pressures of the individual gases present. By successive approximations we can compute the molal volume V of the mixture from equation 17, which may be written:

$$p = \Phi_1(T, x_1/V) + \Phi_2(T, x_2/V) + \cdots \qquad (34)$$

the desired equilibrium pressures $p_{e1} \equiv \Phi_1(T, x_1/V)$, etc. being evaluated at the same time. The equilibrium pressure of the vapor so determined should agree with the vapor pressure of the univariant system when corrected for the effect of hydrostatic pressure (Poynting relation) and for the solubility (if any) of the inert gas in the condensed complex.

2. Given the total pressure p, the temperature T, and the equilibrium pressures p_{ek} of $m - 1$ of the m substances in the gas phase, to calculate the composition of the gas phase. The equilibrium pressure of the m^{th} substance is computed by difference. From the relations $p_{e1} = \Phi_1(T, n_1/V)$, etc., we can obtain n_1/V, etc., for each gas, whence

$$x_i = \frac{n_i/V}{\Sigma_i(n_i/V)} \tag{35}$$

In some cases the method may require extrapolation of the equation of state or plot of p *vs.* V for the vapor to pressures above the normal vapor pressure of the liquid. Since the effect of hydrostatic pressure on the vapor pressure of a condensed system is small, the extrapolation is in general not severe.

Lurie and Gillespie (134) and Gillespie (72) applied the Gibbs–Dalton law to their measurements on the system $BaCl_2(s)$, $BaCl_2 \cdot 8NH_3(s)$, $NH_3(g) + N_2(g)$ and to those of Larson and Black (109) on $NH_3(l)$, $NH_3(g) + H_2(g) + N_2(g)$. They found that the deviations of the results derived from the Gibbs–Dalton law from the measured values are in the same direction as those derived from the ideal gas law and, up to moderate pressures, roughly one-half as large. At higher pressures there is an indication that Gibbs–Dalton law deviations may pass through zero and change sign. Gerry and Gillespie (60) found that the Gibbs–Dalton law may give results no better than the ideal gas law when the mole fraction of one substance in the gas phase is very small, as in the case of the effect of air, hydrogen, and carbon dioxide on the equilibrium $I_2(s)$, $I_2(g)$ (6, 7, 40).

Cupples (46) applied a method of calculation which is equivalent to the Gibbs–Dalton law to the measurements of Larson and Black mentioned above. Masson and Dolley (138) found that Dalton's law, a consequence of the Gibbs–Dalton law, gave for ethylene–oxygen and ethylene–argon gas mixtures deviations that rose to a maximum of 7 per cent with increasing pressure, then decreased, passed through zero, and became negative. Bartlett, Cupples, and Tremearne (5) found that Amagat's law gave results better than Dalton's law for mixtures of hydrogen and nitrogen.

VII. THE LEWIS AND RANDALL RULE

Lewis and Randall (125) proposed as a basis for the thermodynamic treatment of gas mixtures the rule: The fugacity f_i of a gas in a mixture is equal to the product of its mole fraction x_i into the fugacity f_{Pk} of the pure gas at the total pressure and temperature of the mixture:

$$f_i = f_{Pk}x_i \qquad (i \equiv k = 1, 2, \cdots) \tag{36}$$

The fugacity of a pure gas is defined (123) through the relations:

$$\mu = RT \ln f + \psi(T); \qquad \lim_{p\to 0} \frac{f}{p} = 1 \quad \text{at all temperatures} \tag{37}$$

Consider a pure gas in isothermal equilibrium with a gas mixture through a membrane permeable to that gas alone. The chemical potential of the gas in the mixture is μ_i and that of the pure gas μ_k, where $\mu_i = \mu_k$. According to the Lewis and Randall rule μ_k and hence μ_i can be calculated by substituting into equation 37 the product of the fugacity of the pure gas at the total pressure and temperature of the mixture (in which state it is not in equilibrium with the mixture) times the mole fraction of the gas in the mixture:

$$\mu_i = RT \ln f_{Pk}x_i + \psi_k(T) \qquad (i \equiv k = 1, 2, \cdots) \quad (38)$$

Thus the Lewis and Randall rule contains the necessary idea of equilibrium for a basis of a thermodynamic treatment of gas mixtures.

The following results may be derived from the rule:

1. Subject to the conditions

$$\left.\begin{aligned} T \text{ (each pure gas)} &= T \text{ (mixture)} \\ p \text{ (each pure gas)} &= p \text{ (mixture)} \\ n_k &= n_i \qquad (i \equiv k = 1, 2, \cdots) \end{aligned}\right\} \quad \text{(C)}$$

the following relations hold:

$$\mu_i = \mu_{Pk} + RT \ln x_i \qquad (i \equiv k = 1, 2, \cdots) \quad (39)$$

$$V = \Sigma_k V_{Pk} \quad (40)$$

$$S = \Sigma_k S_{Pk} - \Sigma_i n_i R \ln x_i \quad (41)$$

$$\left.\begin{aligned} &E = \Sigma_k E_{Pk} \qquad H = \Sigma_k H_{Pk} \qquad F_{VT} = \Sigma_k F_{VT_{Pk}} + \Sigma_i n_i RT \ln x_i \\ &F_{pT} = \Sigma_k F_{pT_{Pk}} + \Sigma_i n_i RT \ln x_i \qquad C_V = \Sigma_k C_{V_{Pk}} \qquad C_p = \Sigma_k C_{p_{Pk}} \end{aligned}\right\} \quad (42)$$

where the subscript P denotes that the conditions C apply. Equation 40 is Amagat's law (1, 2).

2. Subject to conditions C the partial molal volume of a gas in a mixture equals the molal volume of the pure gas:

$$\left(\frac{\partial V}{\partial n_i}\right)_{pTn} = \frac{V_{Pk}}{n_k} \qquad (i \equiv k = 1, 2, \cdots) \quad (43)$$

3. The Lewis and Randall rule leads to a complete solution of the problem of the determination of the thermodynamic properties of gas mixtures from measurements on pure gases only.

The quantity μ_{Pk} can be computed from equation 37 by substitution of f_{Pk} in the right-hand side. Comparison with equation 38 yields equation 39. Differentiation of equation 39 subject to conditions C gives:

$$\left(\frac{\partial \mu_i}{\partial p}\right)_{Tn} = \left(\frac{\partial \mu_{Pk}}{\partial p}\right)_T \qquad (i \equiv k = 1, 2, \cdots) \quad (44)$$

$$\left(\frac{\partial \mu_i}{\partial T}\right)_{pn} = \left(\frac{\partial \mu_{Pk}}{\partial T}\right)_p + R \ln x_i \qquad (i \equiv k = 1, 2, \cdots) \quad (45)$$

where the subscript n denotes constant composition. From general thermodynamic theory the left-hand sides of equations 44 and 45 are equal, respectively, to the partial molal volume and the negative partial molal entropy of the gas i in the mixture and the right-hand sides are the corresponding molal quantities for the pure gas. Thus:

$$\left(\frac{\partial V}{\partial n_i}\right)_{pTn} = \frac{V_{Pk}}{n_k} \qquad (i \equiv k = 1, 2, \cdots) \tag{46}$$

$$\left(\frac{\partial S}{\partial n_i}\right)_{pTn} = \frac{S_{Pk}}{n_k} - R \ln x_i \qquad (i \equiv k = 1, 2, \cdots) \tag{47}$$

The first is equation 43. By Euler's theorem:

$$V = \Sigma_i \left[n_i \left(\frac{\partial V}{\partial n_i}\right)_{pTn} \right] = \Sigma_k V_{Pk} \qquad (i \equiv k = 1, 2, \cdots) \tag{48}$$

$$S = \Sigma_i \left[n_i \left(\frac{\partial S}{\partial n_i}\right)_{pTn} \right] = \Sigma_k S_{Pk} - \Sigma_i n_i R \ln x_i \qquad (i \equiv k = 1, 2, \cdots) \tag{49}$$

Thus equations 40 and 41 are proved. Substitution from equations 39, 40, and 41 into the general thermodynamic relations given in Section 2 (page 146) and use of the conditions C give the expressions 42.

Statement 3 follows directly from equations 39 through 42.

VIII. THERMODYNAMIC RELATIONS FOR GAS MIXTURES FROM THE LEWIS AND RANDALL RULE

Let

$$\frac{V}{n_k} = \Phi_k'(p, T) \tag{50}$$

be the equation of state of a pure gas. The Lewis and Randall rule gives the following expression for the volume of a gas mixture:

$$V = \Sigma_k [n_k \Phi_k'(p, T)] \tag{51}$$

where conditions C apply and the summation extends over all of the gases in the mixture. Equation 51 is an equation of state for gas mixtures where the equation for each pure gas is of the form of equation 50. As in the case of the Gibbs–Dalton law, the form of the function Φ_k' may be different for the different gases composing the mixture, and some or all of the functions may be evaluated graphically.

In terms of the thermodynamic properties of pure gases derived by the general limit method, the Lewis and Randall rule gives the following relations for gas mixtures:

$$E = \Sigma_k \left\{ \int_0^p \left[V_k - T \left(\frac{\partial V_k}{\partial T}\right)_{p_k} \right] dp_k - pV_k + n_k h_k^0 \right\} \tag{52}$$

$$H = \Sigma_k \left\{ \int_0^p \left[V_k - T \left(\frac{\partial V_k}{\partial T}\right)_{p_k} \right] dp_k + n_k h_k^0 \right\} \tag{53}$$

$$S = \Sigma_k \left\{ \int_0^p \left[\frac{n_k R}{p_k} - \left(\frac{\partial V_k}{\partial T}\right)_{p_k} \right] dp_k + n_k S_k^0 \right\} - \Sigma_i n_i R \ln p x_i \tag{54}$$

$$F_{VT} = \Sigma_k \left\{ \int_0^p \left[V_k - \frac{n_k RT}{p_k} \right] \mathrm{d}p_k - pV_k + n_k(h_k^0 - Ts_k^0) \right. \qquad\qquad + \Sigma_i n_i RT \ln px_i \tag{55}$$

$$F_{pT} = \Sigma_k \left\{ \int_0^p \left[V_k - \frac{n_k RT}{p_k} \right] \mathrm{d}p_k + n_k(h_k^0 - Ts_k^0) \right\} \qquad\qquad + \Sigma_i n_i RT \ln px_i \tag{56}$$

$$C_V = \Sigma_k \left\{ n_k c_{vk}^* + n_k R + T \frac{(\partial V_k/\partial T)_{p_k}^2}{(\partial V_k/\partial p_k)_T} - T \int_0^p \left(\frac{\partial^2 V_k}{\partial T^2} \right)_{p_k} \mathrm{d}p_k \right\} \tag{57}$$

$$C_p = \Sigma_k \left\{ n_k c_{pk}^* - T \int_0^p \left(\frac{\partial^2 V_k}{\partial T^2} \right)_{p_k} \mathrm{d}p_k \right\} \tag{58}$$

$$\mu_i = \int_0^p \left[\frac{V_k}{n_k} - \frac{RT}{p_k} \right] \mathrm{d}p_k + RT \ln px_i + h_k^0 - Ts_k^0 \tag{59}$$

$$RT \ln f_i = \int_0^p \left[\frac{V_k}{n_k} - \frac{RT}{p_k} \right] \mathrm{d}p_k + RT \ln px_i \tag{60}$$

In the above relations the summations extend over all gases in the mixture, and:

$$h_k^0 = \int_{T_0}^T c_{pk}^* \, \mathrm{d}T + h_{0k}^0 \tag{61}$$

$$s_k^0 = \int_{T_0}^T \frac{c_{pk}^*}{T} \, \mathrm{d}T + s_{0k}^0 \tag{62}$$

The quantities h_{0k}^0 and s_{0k}^0 are the heat content and entropy constants for the pure gas k at the temperature T_0; c_{pk}^* is the molal constant-pressure heat capacity at zero pressure.

The left-hand sides of equations 52 through 58 refer to the gas mixture at p, T, and the mole numbers n_i ($i \equiv k = 1, 2, \cdots$). The terms enclosed in braces are to be computed for each pure gas k from equations 50, 61, and 62 with p, T, and n_k having the same value for each pure gas as in the mixture, and the sum taken for all gases composing the mixture. The terms not enclosed in braces are to be computed for the gas mixture. In equations 59 and 60 the left-hand side refers to a gas in a mixture, and the right-hand side is to be computed as just described.

The Lewis and Randall rule leads directly to an expression for the mass action constant, K_p. Let K_p^* be the value of K_p at zero pressure. Substitution from equation 59 into the equilibrium expression $\Sigma_i \nu_i \mu_i = 0$ gives:

$$\ln \frac{K_p}{K_p^*} = -\frac{1}{RT} \Sigma_k \left\{ \nu_k \left[\int_0^p \left[\frac{V_k}{n_k} - \frac{RT}{p_k} \right] \mathrm{d}p_k \right] \right\} \tag{63}$$

$$\ln K_p^* = -\frac{1}{RT} \Sigma_k \{ \nu_k [h_k^0 - Ts_k^0] \} \tag{64}$$

Since the integrand of equation 63 is a function of p and T alone, the expression for the mass action function given by the Lewis and Randall rule is independent of the composition of the equilibrium mixture.

IX. APPLICATION OF THE LEWIS AND RANDALL RULE TO MIXTURES OF REAL GASES

A. Compressibility of gas mixtures

In a later section there will be derived a thermodynamic relation for the fugacity f_i of a gas i at a mole fraction x_i in a gas mixture at total pressure p and temperature T:

$$\ln f_i = \frac{1}{RT}\int_0^p \left[\left(\frac{\partial V}{\partial n_i}\right)_{pTn} - \frac{RT}{p}\right] dp + \ln px_i \tag{65}$$

where $(\partial V/\partial n_i)_{pTn}$ is the partial molal volume of the gas in the given mixture. The thermodynamic relation for the fugacity f_k of a pure gas at pressure p and temperature T is:

$$\ln f_k = \frac{1}{RT}\int_0^p \left[\frac{V_k}{n_k} - \frac{RT}{p_k}\right] dp_k + \ln p \tag{66}$$

where V/n_k is the molal volume of the pure gas. The Lewis and Randall rule, $f_i = f_{Pk}x_i$, yields directly equation 60, where f_{Pk} is computed from equation 66 for the total pressure and temperature of the mixture. The error of the Lewis and Randall rule expressed as a ratio is, from equations 60 and 65:

$$\ln \frac{f_i}{f_{Pk}\,x_i} = \frac{1}{RT}\int_0^p \left[\left(\frac{\partial V}{\partial n_i}\right)_{pTn} - \frac{V_k}{n_k}\right] dp \tag{67}$$

The integrand is the increase in volume of a system consisting of a very great quantity of the mixture together with a quantity of the pure gas at the same temperature and pressure when 1 mole of gas is removed from the pure gas and added to the mixture at constant temperature and pressure. It was called the expansion of the gas on mixing by Gillespie (68, 69), who studied this quantity for argon–ethylene (138), hydrogen–nitrogen (4), and methane–nitrogen (95) mixtures. In general the integrand of equation 67 is not zero, as would be required by the Lewis and Randall rule, nor is it zero at zero pressure. The first mixture deviates rather widely from the rule; the other two give much smaller expansions.

Gibson and Sosnick (63), Randall and Sosnick (147), Merz and Whittaker (140), and Krichevsky (97) studied the deviations of the above three mixtures from the Lewis and Randall rule and Sage, Lacey, and their collaborators (151, 152, 154, 157) studied the deviations in mixtures of methane with ethane, n-butane, and n-pentane. Amagat's law, a consequence of the Lewis and Randall rule, was applied by Masson and Dolley (138) to ethylene–argon, ethylene–oxygen, and argon–oxygen mixtures and by Bartlett, Cupples, and Tremearne (5) to a hydrogen–nitrogen mixture.

The results of these comparisons may be summarized as follows: (*1*) when all gases in the mixture are far above their critical temperatures the Lewis and Randall rule gives good results up to 100 atm. and fair results to several hundred atmospheres; (*2*) when one of the gases in the mixture is close to or below its critical point the rule may fail badly even at 50 atm.; (*3*) for paraffin hydrocarbon mixtures up to 200 atm. the rule gives results to 10 or 15 per cent.

B. Effect of an inert gas on univariant equilibria

Lurie and Gillespie (134) applied the Lewis and Randall rule to their measurements of total pressures and gas compositions in the system $BaCl_2(s)$, $BaCl_2 \cdot 8NH_3(s)$, $NH_3(g) + N_2(g)$ at 45°C. At this temperature the pressure of ammonia above the condensed complex is 7 atm. and that of liquid ammonia is about 17 atm. Strictly, the application of the rule to ammonia is limited to the vapor pressure of the liquid, since in equation 59 the upper limit of the integral is the total pressure in the gas phase. In the region in which the rule could be applied without extrapolation, it gave a fair representation of the observed equilibrium pressures of ammonia.

The calculations were made as follows:

1. Given the total pressure, temperature, and composition of the gas phase, to determine the equilibrium pressures of the constituent gases. We can compute by equation 60 the fugacity of gas i in the mixture. This is the fugacity f_k on the left-hand side of equation 66, so that the corresponding equilibrium pressure p_e of the gas can be calculated from equation 66 by successive approximations.

2. Given the total pressure p, temperature T, and equilibrium pressures $p_{e1}, p_{e2} \cdots$ of $m - 1$ of the m substances present in the gas mixture, to find the composition of the gas phase. We can calculate by equation 66 the corresponding fugacities of the $m - 1$ pure gases at T and p_{e1}, p_{e2}, etc. Substitution of each fugacity into equation 60 gives the corresponding mole fraction of the gas.

Randall and Sosnick (147) extended this calculation to apply to the entire pressure range of Lurie and Gillespie's measurements, to the effects of nitrogen and hydrogen on the equilibrium $CO_2(l) = CO_2(g)$ (145), and to the effects of hydrogen, nitrogen, and their mixtures on the equilibrium $H_2O(l) = H_2O(g)$ (3). This requires an extrapolation of the equation of state to pressure far above the vapor pressure of the liquid, but Randall and Sosnick conclude that such an extrapolation is permissible to the maximum on the James Thomson continuous isotherm. It is difficult to determine from such calculations what part of the disagreement is due to failure of the equation of state extrapolation.

Gerry (quoted in reference 73) found the Lewis and Randall rule "extremely bad for cases of condensible substances at small mole fraction." The system studied was $I_2(s)$, $I_2(g) + X(g)$, where X is air, hydrogen, or carbon dioxide, and the total pressure 1 atm. The errors of the rule greatly exceeded those of the ordinary ideal gas treatment.

C. Chemical equilibria

The Lewis and Randall rule is especially useful in computing the effect of pressure on chemical equilibria in gaseous systems. Given the value of K_p^*, which is a function of the temperature alone, and an equation of state with known values of the constants for each of the pure gases or an equivalent graphical method for evaluating the integrals of equation 63, we can compute the value of K_p at any pressure. From this and the composition of the gas mixture before chemical reaction has occurred we can obtain the composition of the equilibrium mixture. The advantage of this method is that the composition of the equilibrium mixture does not occur on the right-hand side of the equation.

Gillespie (67) first derived equation 63 and applied it to the measurements of the equilibrium $3H_2 + N_2 = 2NH_3$ by Larson and Dodge (108, 110). He evaluated the integrals from the Keyes equation of state (90). The agreement between observed and calculated values of K_p was good up to 300 atm. Von Wettberg and Dodge (182) applied the same method to the equilibrium $CO + 2H_2 = CH_3OH$, evaluating the integrals from the van der Waals equation. The rather wide scattering of some of the points was attributed to lack of equilibrium.

A second method of applying the Lewis and Randall rule to chemical equilibria in gaseous systems differs from the preceding only in the arrangement of the calculations to suit graphical or tabular representation of fugacities. Applying the equilibrium relation $\Sigma_i \nu_i \mu_i = 0$, we obtain from equations 59 and 60 for a chemical reaction at equilibrium:

$$\Sigma_i \nu_i \ln f_i \equiv \ln K_f = -\frac{1}{RT} \Sigma_k \{\nu_k [h_k^0 - Ts_k^0]\} \tag{68}$$

where f_i is the actual fugacity of gas i in the equilibrium mixture. The Lewis and Randall rule gives $f_i = f_{Pk} x_i = (f_{Pk}/p) p x_i$, where p is the pressure of the equilibrium mixture. Thus:

$$\Sigma_i \nu_i \ln f_i = \Sigma_k \nu_k \ln \frac{f_{Pk}}{p} + \Sigma_i \nu_i \ln p x_i \tag{69}$$

or

$$\ln \frac{K_p}{K_f} = \ln \frac{K_p}{K_p^*} = -\Sigma_k \nu_k \ln \frac{f_{Pk}}{p} \tag{70}$$

Values of (f_{Pk}/p) can be computed by use of equation 66 from isothermal compressibility data. Tables of this quantity at various temperatures and pressures have been prepared by Deming and Shupe (49, 51, 52, 53) for a number of permanent gases; by Sage, Lacey, and their collaborators for many hydrocarbons (153, 155, 156, 158, 159); and by other investigators (43, 54, 89). Given K_p^* and values of f_{Pk}/p for the reacting gases we can compute K_p and hence the equilibrium composition of the mixture.

Brown and his collaborators (171) presented on a single plot curves of (f/p) against p_r ($= p/p_{\text{critical}}$) for a number of different values of T_r ($= T/T_{\text{critical}}$) for methane and for pentane. Lewis and his collaborators (126, 127, 128, 129, 130, 131) suggested that for paraffin hydrocarbons of more than two carbon atoms (f/p) can be considered as depending only on p_r and T_r independent of the hydrocarbon (law of corresponding states). They therefore presented a generalized plot of (f_{Pk}/p) *vs.* p_r for various values of T_r good for all such hydrocarbons. Newton (141) presented such a plot not restricted to hydrocarbons, and Newton and Dodge (142, 143) applied equation 70 to the ammonia and methanol synthesis equilibria, evaluating (f_{Pk}/p) from the plot and the critical constants of the gases involved. The agreement between observed and calculated values of K_p was good to 300 atm., fair at 600 atm., but poor at 1000 atm., a result which they attributed to failure of the Lewis and Randall rule.

D. Liquid–vapor equilibria

Consider the equilibrium between the liquid and vapor phases of a binary or multicomponent system at the total pressure p and temperature T. If the liquid is a perfect solution (124) the fugacity of a substance i in the liquid phase is $f_{iL}y_i$, where f_{iL} is the fugacity of the pure liquid of substance i at p and T and y_i its mole fraction in the solution. The Lewis and Randall rule gives for the fugacity of i in the vapor $f_{Pk}x_i$, where f_{Pk} is the fugacity of the pure vapor of i at p and T, and x_i its mole fraction in the vapor. At equilibrium:

$$f_{iL}y_i = f_{Pk}x_i \qquad (i \equiv k = 1, 2, \cdots) \quad (71)$$

Brown and his collaborators (171, 173) discussed the application of this equation to the design of high-pressure hydrocarbon equipment. In general the evaluation of f_{Pk} from equation 66 or from fugacity charts requires extrapolation to pressures above the condensation pressure of the vapor.

Lewis and Luke (130, 131) made the approximation that f_{iL} may be replaced by the fugacity of the pure liquid at its vapor pressure and T. This is equal to the fugacity of the pure vapor at the vapor pressure and T and can be read from the generalized fugacity chart if the vapor pressure at T and the critical constants of the substance are known.

Although the approximations are more severe than those of the Lewis and Randall rule, the results are sufficiently accurate for the design of high-pressure hydrocarbon equipment.

Dodge and Newton (55) derived for liquid–vapor equilibria relations that involve the use of the Lewis and Randall rule for the vapor phase but not the use of generalized fugacity charts.

X. COMPARISON OF THE GIBBS–DALTON LAW AND THE LEWIS AND RANDALL RULE

The Gibbs–Dalton law requires Dalton's law of the additivity of pressure at constant volume and temperature (equation 6) to hold; it leads to the additivity of certain thermodynamic quantities at constant volume and temperature (equation 8); and it requires that $(\partial p/\partial n_i)_{VTn}$ for the mixture equal $(\partial p/\partial n)_{VT}$ for each pure gas.

The Lewis and Randall rule yields Amagat's law of the additivity of volumes at constant total pressure and temperature (equation 40); it gives the result that the difference between certain thermodynamic properties of a gas mixture and the sum of the corresponding properties for the constituents at the total pressure and temperature of the mixture is the same as for ideal gases (equations 39, 41, and 42); and it requires that the partial molal volume $(\partial V/\partial n_i)_{pTn}$ of a gas in a mixture equal its molal volume V/n.

In general, Amagat's law gives results superior to Dalton's law for gas mixtures at temperatures well above the critical temperature of each constituent, while Dalton's law seems to give the better results if one of the constituent gases is below or only slightly above the critical temperature.

The Gibbs–Dalton law is readily applied to the calculation of the composition of the gas phase in equilibrium with a condensed complex yielding one of the constituents. It requires a much shorter extrapolation of the compressibility data on the condensible vapor above the condensation pressure than the Lewis and Randall rule. Gillespie (72) states that at not too high pressures the Gibbs–Dalton law yields too low and the Lewis and Randall rule too high a mole fraction for the vapor that is in equilibrium with the condensed complex.

The Lewis and Randall rule yields a simple expression for the mass action function and the application of the equation to chemical equilibria is facilitated by the use of equation 70 and generalized fugacity charts, although the latter introduces the further approximation of the law of corresponding states. The results are satisfactory to relatively high pressures. The Gibbs–Dalton expression is more involved and has never been applied.

The use of the Lewis and Randall rule together with generalized fugacity charts for the vapor phase of a multicomponent liquid–vapor equilibrium seems to give results sufficiently accurate for engineering calculations. It is possible that the Gibbs–Dalton law would yield more accurate results, but the work would be increased.

XI. THE GENERAL LIMIT METHOD

The procedure called by Gillespie (73) the general limit method assumes that, as the pressure on a real gas or gaseous mixture is reduced at constant temperature and mole numbers, certain thermodynamic properties of a system of real gases approach those of ideal gases in a definite manner. The method was applied by van der Waals (181) to derive an expression for the work content (F_{VT}) of a gas mixture but without elimination of an improper integral introduced through the equation for the entropy. van Laar and Lorenz (106, 107) and Keyes (92) employed the method for the energy of a gas mixture where improper integrals do not appear.

Gillespie (66) was the first to apply the method to the calculation of the fugacity of a gas in a mixture and to derive a general equation for the effect of pressure and composition on chemical equilibrium. He showed how the improper integral that had given difficulty in the past could be transformed by the addition and subtraction of a term. He integrated the relations by means of an equation of state for gas mixtures and made extensive application of the resulting equations to the effect of pressure on chemical equilibrium (67) and the calculation of the equilibrium pressure of a gas in a gaseous mixture (60, 68, 134). The equations gave a good representation of the experimental data.

de Donder (56, 57, 149) used the general limit method to write down directly the work content (F_{VT}) of a gaseous mixture. The method has been applied by van Lerberghe (118, 119, 120, 121), and by Prigogine and Defay (146) to the calculation of the fugacity of a gas in a mixture.

Equations for the fugacity of a gas in a mixture equivalent to that derived by Gillespie were obtained by Gibson and Sosnick (63), Merz and Whittaker (140), van Lerberghe (120), Tunnell (180), Goranson (83), and Sage and Lacey (154).

It has been found (15) that the thermodynamic properties of real gases and gas mixtures, expressed in the form of definite integrals, can be derived from two assumptions: one concerning the isothermal variation of the energy of pure gases at low pressure and the other concerning the isothermal variation of the ratio of the equilibrium pressure of a gas in a mixture to its partial pressure at low pressures. The resulting expressions (8, 9, 10) may be considered exact and state the thermodynamic functions of a gas mixture in terms of: (*1*) definite integrals that depend on the properties of the gas mixture for the effect of pressure; (*2*) definite integrals that depend on the heat capacities of the pure constituent gases at zero pressure for the temperature effect; (*3*) certain integration constants for the pure constituent gases; and (*4*) the temperature, pressure (or density), and composition of the mixture. The completion of the solution of the problem is attained on evaluating the integrals of (*1*) above by an equation of state for gas mixtures in which the parameters of the mixture are computed from those of the pure gases and the composition from certain relations usually called combination of constants. The final equations are approximate to the extent that the equation of state for mixtures does not represent the variation of the volumetric behavior of gas mixtures with pressure, temperature, and composition. The resulting equation (75, 76, 77) expressing the mass action constant of a gaseous reaction as a function of pressure, temperature, and composition of the equilibrium mixture contains, in principle, no arbitrary constants, all parameters being evaluated from the volumetric and thermal properties of the pure constituent gases. The equation has been applied to the equilibrium measurements of Larson and Dodge (108, 110) and of Haber (84) for the ammonia synthesis reaction with two of the parameters considered adjustable (those for ΔH and ΔS of the chemical reaction in the standard state). For the first time the three sets of measurements are correlated in a satisfactory method.

Scatchard and Raymond (165) and Scatchard and coworkers (164, 166, 167, 168, 169, 170) have used the relations resulting from the general limit method for the gas phase in the derivation of equations for liquid–vapor equilibria.

Benedict, Webb, and Rubin (38) applied the general limit method to the work content of a gas mixture and employed the results of Mayer (139) to prove that the limit of the quantity they used exists as the pressure approaches zero. The resulting relations were integrated by use of their equation of state (37) together with certain rules for combination of constants. They applied the relations with excellent results to binary (38) and ternary (35, 36) equilibrium mixtures of hydrocarbons in the liquid and gaseous phases.

Although different investigators applied the general limit method to different thermodynamic functions, all arrived at the same equation for the effect of pressure and composition on the fugacity (or, what amounts to the same thing, the chemical potential) of a gas in a mixture—the equation obtained by Gillespie.

The complete solution of the problem of expressing the properties of gas mixtures in terms of those of the pure constituent gases thus involves two steps: (*1*) a general limit method for the exact equations; and (*2*) combination of constants of an equation of state for evaluating certain definite integrals of (*1*).

XII. THE GENERAL LIMIT METHOD FOR PURE GASES

A. The assumptions

Although two assumptions suffice for deriving general relations for the thermodynamic properties of pure gases and gas mixtures, it is desirable to divorce the treatment for gas mixtures from that of pure gases. If this is done we need three assumptions: two for pure gases and of course one additional for gas mixtures.

We shall consider two pressure regions: (*1*) low pressures, at which terms of the order of p^2 are negligible in comparison with terms of the order of p or of unity; and (*2*) very low pressures, at which terms of the order of p are negligible in comparison with unity. Thermodynamic quantities in the region of very low pressure will be denoted by a superscript asterisk.

The two assumptions for pure gases are:

1. At all temperatures[2] the pressure–volume product of a fixed mass of a real gas can be represented at constant temperature and in the region of low pressure by the expression:

$$pV = F(T,m) + A(T,m)p + O(p^2) \qquad (T,m \text{ constant}) \qquad (72)$$

where $F(T,m)$ and $A(T,m)$ depend for a given gas on the temperature T and the mass m, $A(T, m)$ and its temperature derivative are bounded, and $O(p^2)$ represents terms of the order of p^2.

2. At all temperatures[2] the heat content of a fixed mass of a real gas can be represented at constant temperature and in the region of low pressure by the expression:

$$H = f(T,m) + B(T,m)p + O(p^2) \qquad (T,m \text{ constant}) \qquad (73)$$

where $f(T,m)$ and $B(T,m)$ depend for a given gas on T and m, and B and its temperature derivative are bounded.

Substitution from equations 72 and 73 into the general relation:

$$\left(\frac{\partial H}{\partial p}\right)_T = V - T\left(\frac{\partial V}{\partial T}\right)_p \qquad (74)$$

yields

$$B + O(p) = \frac{1}{p}\left[F - T\left(\frac{\partial F}{\partial T}\right)_m\right] + \left[A - T\left(\frac{\partial A}{\partial T}\right)_m\right] + O(p) \qquad (75)$$

for a fixed mass of a gas in the region of low pressure. Equation 75 can hold only if

$$F - T\left(\frac{\partial F}{\partial T}\right)_m = 0$$

---

[2] The region in the neighborhood of 0°K. is to be excluded.

Integration at constant m gives

$$F = r(m)T \tag{76}$$

where T is thermodynamic temperature and $r(m)$ depends only on m for a given gas. Thus equation 72 becomes:

$$pV = r(m)T + A(T,m)p + O(p^2) \qquad (m \text{ constant}) \tag{77}$$

and in the region of very low pressures:

$$\frac{(pV)^*}{rT} = 1 \tag{78}$$

Since volume is an extensive property of a system, the value of r will be characteristic of each gas if we replace V in equation 78 by V/m, where m is the number of units of mass of the gas in the volume V. Further let m be expressed in terms of moles where 1 mole of a gas is that mass which at a very low pressure and each temperature occupies the same volume as 32 g. of gaseous oxygen measured at the same pressure and temperature. If m is so measured, r is independent of the gas and will be denoted by R. Thus for low pressures we may write:

$$pV = nRT + \frac{n\beta p}{RT} + O(p^2) \tag{79}$$

$$pV = nRT + \frac{n^2\beta}{V} + O\left(\frac{n^2}{V^2}\right) \tag{80}$$

where n is the number of moles in the volume V, β is a pure temperature function ($= ART$ if A is taken for 1 mole of gas) and is so chosen to conform to the usual writing of equations of state.

B. The integration path

We can now proceed to write down all of the thermodynamic functions of a pure gas in terms of the independent variables p and T, or V and T, and transform the improper integrals that occur in the expression for the entropy.

The reference state of a gas is taken at any arbitrarily chosen pressure (p_0) and temperature (T_0), the corresponding volume of n moles being V_0. Since the thermodynamic functions in which we are interested are all state variables we may proceed from the standard state (p_0, T_0) or (V_0, T_0) to any other state (p, T) or (V, T) along any path. We shall use a three-step path involving a state in the region of very low pressure p^* (or V^*). Let the mass of gas remain constant and:

1. Change the pressure from p_0 to p^* (or the volume from the corresponding values V_0 to V^*) at the temperature T_0.
2. Change the temperature from T_0 to T with the pressure held constant at p^* (or the volume at V^*).
3. Change the pressure from p^* to p (or the volume from V^* to V) at the temperature T.

C. Derivations for the independent variables p and T

From the usual thermodynamic relations we find for n moles of a real gas:

$$H(p,T) - H(p_0,T_0) = {}_{T_0}\int_{p_0}^{p_*} \left[V - T\left(\frac{\partial V}{\partial T}\right)_p\right] \mathrm{d}p + {}_{p_*}\int_{T_0}^{T} nc_p^* \,\mathrm{d}T + {}_{T}\int_{p_*}^{p} \left[V - T\left(\frac{\partial V}{\partial T}\right)_p\right] \mathrm{d}p \quad (81)$$

$$S(p,T) - S(p_0,T_0) = -{}_{T_0}\int_{p_0}^{p_*} \left(\frac{\partial V}{\partial T}\right)_p \mathrm{d}p + {}_{p_*}\int_{T_0}^{T} \frac{nc_p^*}{T} \mathrm{d}T - {}_{T}\int_{p_*}^{p} \left(\frac{\partial V}{\partial T}\right)_p \mathrm{d}p \quad (82)$$

$$= {}_{T_0}\int_{p_0}^{p_*} \left[\frac{nR}{p} - \left(\frac{\partial V}{\partial T}\right)_p\right] \mathrm{d}p + {}_{p_*}\int_{T_0}^{T} \frac{nc_p^*}{T} \mathrm{d}T + {}_{T}\int_{p_*}^{p} \left[\frac{nR}{p} - \left(\frac{\partial V}{\partial T}\right)_p\right] \mathrm{d}p - nR \ln \frac{p}{p_0} \quad (83)$$

We can make the following simplifications:

1. Let:

$$h^0 = h_0^0 + \int_{T_0}^{T} c_p^* \,\mathrm{d}T; \qquad nh_0^0 = H(p_0,T_0) + {}_{T_0}\int_{p_0}^{p_*} \left[V - T\left(\frac{\partial V}{\partial T}\right)_p\right] \mathrm{d}p \quad (84)$$

$$s^0 = s_0^0 + \int_{T_0}^{T} \frac{c_p^*}{T} \mathrm{d}T; \qquad ns_0^0 = S(p_0, T_0) + {}_{T_0}\int_{p_0}^{p_*} \left[\frac{nR}{p} - \left(\frac{\partial V}{\partial T}\right)_p\right] \mathrm{d}p + nR \ln p_0 \quad (85)$$

2. From equation 73 and the relation $C_p = (\partial H/\partial T)_p$ we find that

$$C_p^* = \left(\frac{\partial f}{\partial T}\right)_m + O(p) \quad (86)$$

and hence c_p^* is a function of T alone.

3. From equation 79 we find:

$$\lim_{p_* \to 0} \int_{p_*}^{0} \left[V - T\left(\frac{\partial V}{\partial T}\right)_p\right] \mathrm{d}p = \lim_{p_* \to 0} \int_{p_*}^{0} \left[\frac{nR}{p} - \left(\frac{\partial V}{\partial T}\right)_p\right] \mathrm{d}p = 0 \quad (87)$$

Equations 81 and 83 can now be written:

$$H(p, T) = \int_0^p \left[V - T\left(\frac{\partial V}{\partial T}\right)_p\right] \mathrm{d}p + nh^0 \quad (88)$$

$$S(p, T) = \int_0^p \left[\frac{nR}{p} - \left(\frac{\partial V}{\partial T}\right)_p\right] \mathrm{d}p - nR \ln p + ns^0 \quad (89)$$

It is understood that the integrals are to be evaluated at the running temperature T. From equations 88 and 89 the other thermodynamic properties of real gases can be written down by use of the relations given in Section II (page 143). These are collected in Section XII E (page 165).

D. Derivations for the independent variables V and T

In terms of the independent variables V and T we find for n moles of a real gas:,

$$E(V,T) - E(V_0,T_0) = {}_{T_0}\!\int_{V_0}^{V_*}\left[T\left(\frac{\partial p}{\partial T}\right)_V - p\right]\mathrm{d}V + {}_{V_*}\!\int_{T_0}^{T} nc_v^*\,\mathrm{d}T + {}_{T}\!\int_{V_*}^{V}\left[T\left(\frac{\partial p}{\partial T}\right)_V - p\right]\mathrm{d}V \qquad (90)$$

$$S(V,T) - S(V_0,T_0) = {}_{T_0}\!\int_{V_0}^{V_*}\left(\frac{\partial p}{\partial T}\right)_V\mathrm{d}V + {}_{V_*}\!\int_{T_0}^{T}\frac{nc_v^*}{T}\,\mathrm{d}T + {}_{T}\!\int_{V_*}^{V}\left(\frac{\partial p}{\partial T}\right)_V\mathrm{d}V \qquad (91)$$

$$= {}_{T_0}\!\int_{V_0}^{V_*}\left[\left(\frac{\partial p}{\partial T}\right)_V - \frac{nR}{V}\right]\mathrm{d}V + {}_{V_*}\!\int_{T_0}^{T}\frac{nc_v^*}{T}\,\mathrm{d}T + {}_{T}\!\int_{V_*}^{V}\left[\left(\frac{\partial p}{\partial T}\right)_V - \frac{nR}{V}\right]\mathrm{d}V + nR\ln\frac{V}{V_0} \qquad (92)$$

These equations can also be simplified:

1. Let:

$$e^0 = e_0^0 + \int_{T_0}^{T} c_v^*\,\mathrm{d}T; \qquad ne_0^0 = E(V_0,T_0) + {}_{T_0}\!\int_{V_0}^{V_*}\left[T\left(\frac{\partial p}{\partial T}\right)_V - p\right]\mathrm{d}V \qquad (93)$$

Now from equations 84 and 93 we find:

$$nh_0^0 - ne_0^0 = H(p_0,T_0) - E(V_0,T_0) + {}_{T_0}\!\int_{p_0}^{p_*}\left[V - T\left(\frac{\partial V}{\partial T}\right)_p\right]\mathrm{d}p - {}_{T_0}\!\int_{V_0}^{V_*}\left[T\left(\frac{\partial p}{\partial T}\right)_V - p\right]\mathrm{d}V \qquad (94)$$

$$= p_0V_0 + {}_{T_0}\!\int_{p_0V_0}^{p_*V_*}\mathrm{d}(pV)$$

$$= nRT_0$$

Since $S(V_0,T_0) = S(p_0,T_0)$ we can replace the former by the expression for the latter obtained from the second of equations 85. This gives on simplification:

$$S(V_0,T_0) + {}_{T_0}\!\int_{V_0}^{V_*}\left[\left(\frac{\partial p}{\partial T}\right)_V - \frac{nR}{V}\right]\mathrm{d}V - nR\ln V_0$$

$$= -nR\ln p_0V_0 - nR\;{}_{T_0}\!\int_{p_0V_0}^{p_*V_*}\mathrm{d}\ln(pV) + ns_0^0$$

$$= -nR\ln nRT_0 + ns_0^0 \qquad (95)$$

Furthermore, from equation 80 and general thermodynamic relations we find

$$c_p^* - c_v^* = R \qquad (96)$$

and from the first of equation 85:

$$s^0 = s_0^0 + \int_{T_0}^{T} \frac{c_v^*}{T}\,\mathrm{d}T + nR\ln\frac{T}{T_0}$$

2. Use of equation 80 for evaluating the volume integrals shows that V^* can be replaced by infinity.

Thus equations 90 and 92 become:

$$E(V,T) = \int_V^\infty \left[p - T\left(\frac{\partial p}{\partial T}\right)_V\right]\mathrm{d}V + ne^0 \tag{97}$$

$$S(V,T) = \int_V^\infty \left[\frac{nR}{V} - \left(\frac{\partial p}{\partial T}\right)_V\right]\mathrm{d}V + nR\ln\frac{V}{nRT} + ns^0 \tag{98}$$

The relations for the other thermodynamic functions are given in Section XII F (page 166).

E. Equations for pure gases in terms of p and T

$$E = \int_0^p \left[V - T\left(\frac{\partial V}{\partial T}\right)_p\right]\mathrm{d}p - pV + nh^0 \tag{99}$$

$$H = \int_0^p \left[V - T\left(\frac{\partial V}{\partial T}\right)_p\right]\mathrm{d}p + nh^0 \tag{88}$$

$$S = \int_0^p \left[\frac{nR}{p} - \left(\frac{\partial V}{\partial T}\right)_p\right]\mathrm{d}p - nR\ln p + ns^0 \tag{89}$$

$$F_{VT} = \int_0^p \left[V - \frac{nRT}{p}\right]\mathrm{d}p + nRT\ln p - pV + n(h^0 - Ts^0) \tag{100}$$

$$F_{pT} = \int_0^p \left[V - \frac{nRT}{p}\right]\mathrm{d}p + nRT\ln p + n(h^0 - Ts^0) \tag{101}$$

$$\mu = \int_0^p \left[\frac{V}{n} - \frac{RT}{p}\right]\mathrm{d}p + RT\ln p + h^0 - Ts^0 \tag{102}$$

$$RT\ln f = \int_0^p \left[\frac{V}{n} - \frac{RT}{p}\right]\mathrm{d}p + RT\ln p \tag{103}$$

In the above equations the definite integrals are to be evaluated at the temperature T. The integration constants are given by the expressions:

$$h^0 = h_0^0 + \int_{T_0}^{T} c_p^*\,\mathrm{d}T \tag{104}$$

$$s^0 = s_0^0 + \int_{T_0}^{T} \frac{c_p^*}{T}\,\mathrm{d}T \tag{105}$$

where the integrals are to be evaluated along the zero pressure curve and h_0^0 and s_0^0 are the molal heat content and entropy constants for the gas at T_0.

F. Equations for pure gases in terms of V and T

$$E = \int_V^\infty \left[p - T\left(\frac{\partial p}{\partial T}\right)_V\right] \mathrm{d}V + ne^0 \tag{97}$$

$$H = \int_V^\infty \left[p - T\left(\frac{\partial p}{\partial T}\right)_V\right] \mathrm{d}V + pV + ne^0 \tag{106}$$

$$S = \int_V^\infty \left[\frac{nR}{V} - \left(\frac{\partial p}{\partial T}\right)_V\right] \mathrm{d}V + nR \ln \frac{V}{nRT} + ns^0 \tag{98}$$

$$F_{VT} = \int_V^\infty \left[p - \frac{nRT}{V}\right] \mathrm{d}V - nRT \ln \frac{V}{nRT} + n(e^0 - Ts^0) \tag{107}$$

$$F_{pT} = \int_V^\infty \left[p - \frac{nRT}{V}\right] \mathrm{d}V - nRT \ln \frac{V}{nRT} + pV + n(e^0 - Ts^0) \tag{108}$$

$$\mu = \int_V^\infty \left[\left(\frac{\partial p}{\partial n}\right)_{VT} - \frac{RT}{V}\right] \mathrm{d}V - RT \ln \frac{V}{nRT} + RT + e^0 - Ts^0 \tag{109}$$

$$RT \ln f = \int_V^\infty \left[\left(\frac{\partial p}{\partial n}\right)_{VT} - \frac{RT}{V}\right] \mathrm{d}V - RT \ln \frac{V}{nRT} \tag{110}$$

In these relations the definite integrals are to be evaluated at the temperature T, and

$$e^0 = e_0^0 + \int_{T_0}^{T} c_v^* \,\mathrm{d}T; \qquad e_0^0 = h_0^0 - RT_0 \tag{111}$$

$$s^0 = s_0^0 + \int_{T_0}^{T} \frac{c_v^*}{T} \,\mathrm{d}T + R \ln \frac{T}{T_0} \tag{112}$$

The temperature integral is to be evaluated along the zero pressure curve.

XIII. THE GENERAL LIMIT METHOD FOR GAS MIXTURES

A. The assumption

The Gibbs–Dalton law and the Lewis and Randall rule each make one assumption as a basis for the thermodynamic treatment of gas mixture. We shall make the following assumption:

At all temperatures the ratio of the equilibrium pressure of each gas in a mixture to the product of the mole fraction of that gas in the mixture into the total pressure of the mixture can be represented in the region of low pressures by the expression:

$$\frac{p_{ek}}{px_i} = 1 + D_i(T, x_1, x_2, \cdots)p + O(p^2) \tag{113}$$

where D_i depends only on the temperature and composition of the mixture, and D_i and its derivatives with respect to T, n_1, n_2, $\cdots$ are bounded.

B. Relations for the region of low pressure

Consider a system of $\Sigma_i n_i$ moles of a gas mixture and k systems each containing n_{ek} moles of one of the constituent gases. In all variations let:

$$\left.\begin{aligned} T \text{ (each pure gas)} &= T \text{ (mixture)} \\ V \text{ (each pure gas)} &= V \text{ (mixture)} \\ \mu_k &= \mu_i \qquad (k \equiv i = 1, 2, \end{aligned}\right\} \quad \text{(D)}$$

Each pure gas would be in equilibrium with the mixture if separated from it by a membrane permeable to that gas alone. Hence the properties of the pure gases may be called equilibrium properties of the mixture.

Subject to the condition D the general limit method (equation 113) gives the following results for the region of low pressure:

$$1.\ \Sigma_k p_{ek} = p(1 + D_m p) \tag{114}$$

where D_m is a function of temperature and composition $(= \Sigma_i x_i D_i)$.

$$\begin{aligned} 2.\ \frac{n_{ek}}{V} &= \frac{n_i}{V}(1 + 2D_m p) \qquad\qquad (k \equiv i = 1, 2, \cdots) \\ C_{ek} &= C_i(1 + 2D_m p) \\ &= Cx_i(1 + 2D_m p) \\ n_{ek} &= n_i(1 + 2D_m p) \end{aligned} \tag{115}$$

$$3.\ \Sigma_k S_{ek} = S(1 + 2D_m p) \tag{116}$$

$$\begin{aligned} 4.\ pV &= \Sigma_i n_i RT + \frac{(\Sigma_i n_i)\beta_m}{RT}\, p \\ pV &= \Sigma_i n_i RT + \frac{(\Sigma_i n_i)^2 \beta_m}{V} \end{aligned} \tag{117}$$

where β_m is a bounded function of temperature and the composition of the gas mixture.

$$\left.\begin{aligned} 5.\ \Sigma_k E_{ek} &= E(1 + 2D_m p) + D_m pV & \Sigma_k H_{ek} &= H(1 + 2D_m p) \\ \Sigma_k F_{VT_{ek}} &= F_{VT}(1 + 2D_m p) + D_m pV & \Sigma F_{pT_{ek}} &= F_{pT}(1 + 2D_m p) \\ \Sigma_k C_{V_{ek}} &= C_V(1 + D'_m p) & \Sigma C_{p_{ek}} &= C_p(1 + D''_m p) \end{aligned}\right\} \tag{118}$$

where D'_m and D''_m are bounded functions of temperature and the composition of the gas mixture.

To prove (15) these relations we take as independent variables T, μ_1, μ_2, $\cdots$ for the gas mixture and T, μ_k for each pure gas as in equations 10 and 11. Equation 114 follows from equation 113, which may be written for the region of low pressure:

$$p_{ek} = px_i(1 + D_i p) \tag{119}$$

Differentiating equation 114 first with respect to one of the μ's with T and all of the other μ's constant, and then with respect to T with all of the μ's constant, we find for the region of low pressure:

$$\left(\frac{\partial p_{ek}}{\partial \mu_k}\right)_T = \left(\frac{\partial p}{\partial \mu_i}\right)_{T\mu} (1 + 2D_m p) \qquad (k \equiv i = 1, 2, \cdots) \qquad (120)$$

$$\Sigma_k \left(\frac{\partial p_{ek}}{\partial T}\right)_{\mu k} = \left(\frac{\partial p}{\partial T}\right)_\mu (1 + 2D_m p) \qquad (121)$$

Use of equations 12 and 13 gives the first of equations 115 and equation 116. The remainder of the equations 115 follow from the definitions given in Section II (page 143) and the second of the conditions D.

Division of equation 119 by the first of equations 115 gives for the region of low pressure:,

$$\frac{p_{ek} V}{n_{ek}} = \frac{pV}{\Sigma_i n_i} [1 + (D_i - 2D_m)p] \qquad (122)$$

Now the left-hand side of this equation refers to the pure gas existing by itself and hence equation 79 applies. This gives the first of equations 117 and the second follows from the first for the region of low pressure where the terms $O(p^2)$ or $O(1/V^2)$ are negligible.

Use of the thermodynamic relations mentioned in Section II and the results already proved together with the conditions D leads to equations 118.

C. Relations for the region of very low pressure

As the pressure becomes very small at constant temperature and composition the relations given in Sections XII and XIII B show that the functions pV, E, H, C_V, and C_p for a gas mixture approach constant values while S, F_{VT}, and F_{pT} decrease or increase without limit in the same manner as $\pm \ln p$. But

$$\lim_{p \to 0} p \ln p = 0$$

Hence the relations for gas mixtures in the region of very low pressures may be summarized from the results obtained in Section XIII B (page 167). Subject to the conditions D, we have the following:

1. Equilibrium relationships:

$$\left(\frac{p_{ek}}{px_i}\right)^* = 1; \quad \left(\frac{C_{ek}}{C_i}\right)^* = 1; \quad \left(\frac{C_{ek}}{Cx_i}\right)^* = 1;$$
$$\left(\frac{n_{ek}}{n_i}\right)^* = \qquad (k \equiv 1 = 1, 2, \cdots) \qquad (123)$$

2. Additive relationships:

$$\left(\frac{\Sigma_k p_{ek}}{p}\right)^* = 1; \quad \left(\frac{\Sigma_k C_{ek}}{C}\right)^* \quad 1 \qquad (124)$$

$$S^* = \Sigma_k S_{ek}^* \tag{125}$$

$$\begin{aligned} E^* &= \Sigma_k E_k^*; \quad H^* = \Sigma_k H_k^*; \quad F_{VT}^* = \Sigma_k F_{VT_{ek}}^*; \\ F_{pT}^* &= \Sigma_k F_{pT_{ek}}^*; \quad C_V^* = \Sigma_k C_{V_k}^*; \quad C_p^* = \Sigma_k C_{pk}^* \end{aligned} \tag{126}$$

3. The equation of state (independent of the conditions D):

$$(pV)^* = \Sigma_i n_i RT \tag{127}$$

The designation e can be omitted in the relations for E, H, C_V, and C_p, since these quantities are essentially constant in the region of low pressure at constant temperature and mole numbers.

D. General relations

Corresponding to equations 81 and 83 for the heat content and entropy of a pure gas, we have for $\Sigma_i n_i$ moles of a gas mixture of constant composition:

$$\begin{aligned} H(p,T) - H(p_0,T_0) = {}_{T_0}\!\int_{p_0}^{p^*} \left[V - T\left(\frac{\partial V}{\partial T}\right)_{pn}\right] \mathrm{d}p \\ + {}_{p^*}\!\int_{T_0}^{T} C_p^* \, \mathrm{d}T + {}_{T}\!\int_{p^*}^{p} \left[V - T\left(\frac{\partial V}{\partial T}\right)_{pn}\right] \mathrm{d}p \end{aligned} \tag{128}$$

$$\begin{aligned} S(p,T) - S(p_0,T_0) = {}_{T_0}\!\int_{p_0}^{p^*} \left[\frac{\Sigma_i n_i R}{p} - \left(\frac{\partial V}{\partial T}\right)_{pn}\right] \mathrm{d}p + {}_{p^*}\!\int_{T_0}^{T} \frac{C_p^*}{T} \, \mathrm{d}T \\ + {}_{T}\!\int_{p^*}^{p} \left[\frac{\Sigma_i n_i R}{p} - \left(\frac{\partial V}{\partial T}\right)_{pn}\right] \mathrm{d}p - (\Sigma_i n_i) R \ln \frac{p}{p_0} \end{aligned} \tag{129}$$

where the subscript n denotes constant mole numbers and C_p^* is the heat capacity of $\Sigma_i n_i$ moles of gas mixture at a very low pressure.

Now from equations 126, 88, and 123:

$$H(p^*,T_0) = \Sigma_k H_k(p^*,T_0) = \Sigma_k n_{ek} h_{0k}^0 = \Sigma_i n_i h_{0i}^0 \tag{130}$$

where for convenience h_{0i}^0 is written for h_{0k}^0, since this will lead to no confusion. And from equations 125, 89, and 123:

$$\begin{aligned} S(p^*,T_0) &= \Sigma_k S_{ek}(p_{ek}^*,T_0) = \Sigma_k n_{ek} s_{0k}^0 - \Sigma_k n_{ek} R \ln p_{ek}^* \\ &= \Sigma_i n_i s_{0i}^0 - \Sigma_i (n_i R \ln p^* x_i) \end{aligned} \tag{131}$$

Also from equation 126 and equation 123:

$$C_p^* = \Sigma_k C_{pk}^* = \Sigma_k n_{ek} c_{pk}^* = \Sigma_i n_i c_{pi}^* \tag{132}$$

Thus from equations 128 and 130:

$$H(p_0,T_0) + {}_{T_0}\!\int_{p_0}^{p^*} \left[V - T\left(\frac{\partial V}{\partial T}\right)_{pn}\right] \mathrm{d}p = H(p^*,T_0) = \Sigma_i n_i h_{0i}^0 \tag{133}$$

and from equations 129 and 131:

$$
\begin{aligned}
S(p_0,T_0) + {}_{T_0}\!\int_{p_0}^{p^*} \left[\frac{\Sigma_i n_i R}{p} - \left(\frac{\partial V}{\partial T}\right)_{pn}\right] \mathrm{d}p + (\Sigma_i n_i)R \ln p_0 \\
= S(p^*,T_0) + (\Sigma_i n_i)R \ln p^* \\
= \Sigma_i n_i s^0_{0i} - \Sigma_i(n_i R \ln x_i) \qquad (134)
\end{aligned}
$$

We can now substitute from equations 132, 133, and 134 into equations 128 and 129. Then use of equation 117 shows that we can replace the limit p^* by 0 (compare equation 87), and obtain:

$$H(p,T) = \int_0^p \left[V - T\left(\frac{\partial V}{\partial T}\right)_{pn}\right] \mathrm{d}p + \Sigma_i(n_i h^0_i) \qquad (135)$$

$$S(p,T) = \int_0^p \left[\frac{\Sigma_i n_i R}{p} - \left(\frac{\partial V}{\partial T}\right)_{pn}\right] \mathrm{d}p - \Sigma_i(n_i R \ln p x_i) + \Sigma_i(n_i s^0_i) \qquad (136)$$

where h^0_i and s^0_i are given by equations 104 and 105.

The other thermodynamic properties can now be written down as was done for the pure gases. These are collected in Section XIII E below.

We can proceed in a similar fashion to obtain equations for the energy and entropy of a gas mixture in terms of the independent variables V and T, analogous to equations 97 and 98 for a pure gas. These equations are given in Section XIII F.

E. Equations for gas mixtures in terms of p and T

$$E = \int_0^p \left[V - T\left(\frac{\partial V}{\partial T}\right)_{pn}\right] \mathrm{d}p - pV + \Sigma_i n_i h^0_i \qquad (137)$$

$$H = \int_0^p \left[V - T\left(\frac{\partial V}{\partial T}\right)_{pn}\right] \mathrm{d}p + \Sigma_i n_i h^0_i \qquad (135)$$

$$S = \int_0^p \left[\frac{\Sigma_i n_i R}{p} - \left(\frac{\partial V}{\partial T}\right)_{pn}\right] \mathrm{d}p - \Sigma_i n_i R \ln p x_i + \Sigma_i n_i s^0_i \qquad (136)$$

$$
\begin{aligned}
F_{VT} = \int_0^p \left[V - \frac{\Sigma_i n_i RT}{p}\right] \mathrm{d}p \\
+ \Sigma_i n_i RT \ln p x_i - pV + \Sigma_i n_i(h^0_i - Ts^0_i) \qquad (138)
\end{aligned}
$$

$$F_{pT} = \int_0^p \left[V - \frac{\Sigma_i n_i RT}{p}\right] \mathrm{d}p + \Sigma_i n_i RT \ln p x_i + \Sigma_i n_i(h^0_i - Ts^0_i) \qquad (139)$$

$$\mu_i = \int_0^p \left[\left(\frac{\partial V}{\partial n_i}\right)_{pTn} - \frac{RT}{p}\right] \mathrm{d}p + RT \ln p x_i + h^0_i - Ts^0_i \qquad (140)$$

$$RT \ln f_i = \int_0^p \left[\left(\frac{\partial V}{\partial n_i}\right)_{pTn} - \frac{RT}{p}\right] \mathrm{d}p + RT \ln p x_i \qquad (141)$$

$$h^0_i = h^0_{0i} + \int_{T_0}^T c^*_{pi}\,\mathrm{d}T \qquad s^0_i = s^0_{0i} + \int_{T_0}^T \frac{c^*_{pi}}{T}\,\mathrm{d}T \qquad (142)$$

F. Equations for gas mixtures in terms of V and T

$$E = \int_V^\infty \left[p - T\left(\frac{\partial p}{\partial T}\right)_{Vn}\right] \mathrm{d}V + \Sigma_i n_i e_i^0 \tag{143}$$

$$H = \int_V^\infty \left[p - T\left(\frac{\partial p}{\partial T}\right)_{Vn}\right] \mathrm{d}V + pV + \Sigma_i n_i e_i^0 \tag{144}$$

$$S = \int_V^\infty \left[\frac{\Sigma_i n_i R}{V} - \left(\frac{\partial p}{\partial T}\right)_{Vn}\right] \mathrm{d}V + \Sigma_i n_i R \ln \frac{V}{n_i RT} + \Sigma_i n_i s_i^0 \tag{145}$$

$$F_{VT} = \int_V^\infty \left[p - \frac{\Sigma_i n_i RT}{V}\right] \mathrm{d}V - \Sigma_i n_i RT \ln \frac{V}{n_i RT} + \Sigma_i n_i (e_i^0 - Ts_i^0) \tag{146}$$

$$F_{pT} = \int_V^\infty \left[p - \frac{\Sigma_i n_i RT}{V}\right] \mathrm{d}V - \Sigma_i n_i RT \ln \frac{V}{n_i RT} + pV + \Sigma_i n_i (e_i^0 - Ts_i^0) \tag{147}$$

$$\mu_i = \int_V^\infty \left[\left(\frac{\partial p}{\partial n_i}\right)_{VTn} - \frac{RT}{V}\right] \mathrm{d}V - RT \ln \frac{V}{n_i RT} + RT + e_i^0 - Ts_i^0 \tag{148}$$

$$RT \ln f_i = \int_V^\infty \left[\left(\frac{\partial p}{\partial n_i}\right)_{VTn} - \frac{RT}{V}\right] \mathrm{d}V - RT \ln \frac{V}{n_i RT} \tag{149}$$

$$e_i^0 = e_{0i}^0 + \int_{T_0}^T c_{vi}^* \, \mathrm{d}T \qquad s_i^0 = s_{0i}^0 + \int_{T_0}^T \frac{c_{vi}^*}{T} \mathrm{d}T + R \ln \frac{T}{T_0} \tag{150}$$

G. General equilibrium relationships

1. *Mass action functions for homogeneous gaseous reactions:* Equations 140 and 148 are relations for the chemical potential of a gas in a mixture. Application of the general condition $\Sigma_i \nu_i \mu_i = 0$ gives equations for the mass functions K_p (in terms of p and T) and K_C (in terms of V and T) for a homogeneous gas reaction. To obtain K_p in terms of V and T we add and subtract the term $RT \ln px_i$ to the right-hand side of equation 148 before applying the general condition of equilibrium. These give:

$$\ln \frac{K_p}{K_p^*} = -\frac{1}{RT} \Sigma_i \left\{\nu_i \left[\int_0^p \left[\left(\frac{\partial V}{\partial n_i}\right)_{pTn} - \frac{RT}{p}\right] \mathrm{d}p\right]\right\} \tag{151}$$

$$= -\frac{1}{RT} \Sigma_i \left\{\nu_i \left[\int_V^\infty \left[\left(\frac{\partial p}{\partial n_i}\right)_{VTn} - \frac{RT}{V}\right] \mathrm{d}V\right]\right\} + (\Sigma_i \nu_i) \ln \frac{pV}{(\Sigma_i n_i)RT} \tag{152}$$

$$\ln K_p^* = -\frac{1}{RT} \Sigma_i \left[\nu_i (h_i^0 - Ts_i^0)\right] \tag{153}$$

$$\ln \frac{K_C}{K_C^*} = -\frac{1}{RT} \Sigma_i \left\{\nu_i \left[\int_V^\infty \left[\left(\frac{\partial p}{\partial n_i}\right)_{VTn} - \frac{RT}{p}\right] \mathrm{d}V\right]\right\} \tag{154}$$

$$\ln K_C^* = -\frac{1}{RT} \Sigma_i \left[\nu_i (e_i^0 + RT - Ts_i^0 + RT \ln RT)\right] \tag{155}$$

2. *Equilibrium pressure and equilibrium concentration of a gas in a mixture:* When a pure gas is in isothermal equilibrium through a semipermeable membrane with a mixture containing that gas, the chemical potential of the gas is the same in both phases. Equating the relations for the chemical potential of a pure gas (equations 102 and 109) to those for a gas in a mixture (equations 140 and 148) gives relations for the equilibrium pressure and equilibrium concentration of a gas in a mixture:

$$RT \ln \frac{p_{ek}}{px_i} = \int_0^p \left[\left(\frac{\partial V}{\partial n_i}\right)_{pTn} - \frac{RT}{p}\right] \mathrm{d}p - \int_0^{p_{ek}} \left[\frac{V_k}{n_k} - \frac{RT}{p_k}\right] \mathrm{d}p_k \qquad (156)$$

$$RT \ln \frac{C_{ek}}{C_i} = \int_V^\infty \left[\left(\frac{\partial p}{\partial n_i}\right)_{VTn} - \frac{RT}{V}\right] \mathrm{d}V - \int_V^\infty \left[\left(\frac{\partial p_k}{\partial n_k}\right)_{V_kT} - \frac{RT}{V_k}\right] \mathrm{d}V_k \qquad (157)$$

The first integral on the right-hand side of these equations applies to the gas mixture; the second to the pure gas.

3. *Effect of an inert gas on a univariant equilibrium:* In a univariant equilibrium between a pure substance S_1 in a gas phase and one or more condensed phases:

$$NH_3(l) = NH_3(g) \qquad (p_{01}, T)$$

$$I_2(s) = I_2(g) \qquad (p_{01}, T)$$

$$\frac{1}{8} BaCl_2 \cdot 8NH_3(s) = \frac{1}{8} BaCl_2(s) + NH_3(g) \qquad (p_{01}, T)$$

the pressure p_{01} is a function of temperature alone. The chemical potential μ_{c1} and molal volume $\bar{v}_{c1}$ of S_1 in the condensed phases are:

$$\mu_{c1} = -(\Sigma_j \nu_j \mu_j)/\nu_1; \qquad \bar{v}_{c1} = -(\Sigma_j \nu_j v_j)/\nu_1 \qquad (158)$$

where v_j is the molal volume of the condensed phase j, the algebraic sign of ν_j is negative for a reactant and positive for a product, and the sum extends over all of the condensed phases. In the first two of the above changes in state $\bar{v}_{c1}$ is simply the molal volume of the condensed phase; for the last, Gillespie (71) called $\bar{v}_{c1}$ the partial molal volume of S_1 in the condensed complex.

Let the condensed phases be separated from the gas phase by a semipermeable membrane and subjected to the hydrostatic pressure p whereby the equilibrium pressure of S_1 increases from p_{01} to p_{e1}. Then:

$$\mu_{c1}(p,T) = \mu_{c1}(p_{01},T) + \int_{p_{01}}^{p} \bar{v}_{c1}\, \mathrm{d}p \qquad (159)$$

Now the chemical potentials $\mu_{c1}(p,T)$ and $\mu_{c1}(p_{01},T)$ of S_1 in the condensed complex must equal those of S_1 in the pure gas phase at (p_{e1},T) and (p_{01},T), respectively. These are given by equation 102. Making these substitutions, we obtain:

$$RT \ln \frac{p_{e1}}{p_{01}} = -\int_{p_{01}}^{p_{e1}} \left[\frac{V}{n} - \frac{RT}{p}\right] \mathrm{d}p + \int_{p_{01}}^{p} \bar{v}_{c1}\, \mathrm{d}p \qquad (160)$$

This is the exact Poynting relation.

Now let the univariant system be at equilibrium under its vapor pressure p_{01} and let an inert gas that does not react chemically with S_1 and is insoluble in the condensed complex be introduced into the system until the total pressure becomes p. The equilibrium pressure p_{e1} of S_1 in the system can be computed from equation 156 or from equation 160. These computations will be considered later. If we substitute the value of p_{e1} from equation 160 into equation 156 we obtain:

$$RT \ln \frac{px_1}{p_{01}} = -\int_0^p \left[\left(\frac{\partial V}{\partial n_1}\right)_{pTn} - \frac{RT}{p}\right] dp + \int_0^{p_{01}} \left[\frac{V_k}{n_k} - \frac{RT}{p_k}\right] dp_k + \int_{p_{01}}^{p} \bar{v}_{c1}\, dp \qquad (161)$$

The first integral refers to the gas mixture, the second to the pure substance S_1, and the third to the condensed complex.

4. *Equilibrium between gaseous mixtures and condensed solutions:* The chemical potential $\mu_{ci}(p, T, y_i)$ of a volatile component i of a liquid (or solid) solution at p, T, and mole fraction y_i is:

$$\mu_{ci}(p,T,y_i) = \mu_{ci}(p_{0i},T) + RT \ln y_i + RT \ln \gamma_i + \int_{p_{0i}}^{p} \bar{v}_{ci}\, dp \qquad (162)$$

where $\mu_{ci}\,(p_{01},T)$ is the chemical potential of the component under its own vapor pressure p_{0i} and is given by equation 102, $\bar{v}_{ci}$ is the partial molal volume of the substance S_i in the condensed solution, and γ_i is its activity coefficient in this phase. The latter has been discussed by Scatchard (162, 163), Scatchard and coworkers (164, 165, 166, 167, 168, 169, 170), and Hildebrand (85), who give references to the earlier work. The chemical potential of S_i in the gas phase is given by equation 140. Equating these two expressions we obtain for isothermal equilibrium at T:

$$RT \ln \frac{px_i}{p_{0i}y_i} = -\int_0^p \left[\left(\frac{\partial V}{\partial n_i}\right)_{pTn} - \frac{RT}{p}\right] dp + \int_0^{p_{0i}} \left[\frac{V_k}{n_k} - \frac{RT}{p_k}\right] dp_k + \int_{p_{0i}}^{p} \bar{v}_{ci}\, dp + RT \ln \gamma_i \qquad (i = k) \qquad (163)$$

The first integral is to be evaluated for the gas mixture, the second for the pure gas k, and the third for the condensed phase.

Equation 163 applies directly to each component whose critical temperature is above T. Hildebrand (85) has suggested a method of determining a fictitious value of p_{0i} for a component above its critical temperature.

XIV. INTEGRATION OF THE EQUATIONS

A. The equation of state for pure gases

Aside from graphical methods we can evaluate the definite integrals with respect to pressure or volume occurring in the thermodynamic properties of pure gases from an equation of state of the type $(V/n) = F(p, T)$ or $p = f(V/n, T)$. All of the useful equations of state that have been applied to a number of substances are of the latter type. The Keyes equation (90) has four adjustable constants and gives a fair representation of the volumetric behavior of gases, being especially satisfactory for gases whose isometrics are nearly linear. The equation of Benedict, Webb, and Rubin (37) contains eight adjustable constants and gives an excellent representation of the compressibility data to about twice the critical density. Below the critical temperature this equation gives satisfactory results into the liquid region and permits calculation of the vapor pressure to about 1 per cent. Results of intermediate accuracy are given by the Beattie–Bridgeman equation of state (17, 18, 19, 21) with five adjustable contants. This equation is:

$$p = \frac{nRT}{V} + \frac{n^2\beta}{V^2} + \frac{n^3\gamma}{V^3} + \frac{n^4\delta}{V^4} \tag{164}$$

$$\beta = RTB_0 - A_0 - Rc/T^2$$

$$\gamma = -RTB_0b + A_0a - RB_0c/T^2$$

$$\delta = RB_0bc/T^2$$

The equation has been applied to the compressibility measurements of a number of substances and the values of the constants A_0, a, B_0, b, and c are listed in table 1. The equation gives satisfactory representation of the volumetric behavior of gases to about the critical density and has been used to compute the limiting densities (20), the Joule–Thompson coefficient and pressure variation of the constant pressure heat capacity (12, 13, 14, 41), the heat content and entropy (78, 81, 82, 148), the chemical potential (79), and the fugacities (113, 135, 137) of gases. In general the calculated values agree well with experimental results.

Equation 164 has two main defects. The first is that the second virial coefficient divided by RT, β/RT, approaches the limit B_0 at high temperatures instead of passing through a maximum and then decreasing with increasing temperature. This is more a theoretical than a practical defect: helium is the only gas exhibiting such a maximum below room temperature (at about −75°C.); neon and hydrogen have maxima at about 400°C., and other gases probably well above 1000°C. The second defect has more practical importance. The constant c, which was introduced to account for the curvature of the isometrics, should be a function of density and is so expressed in the equation of Benedict, Webb, and Rubin, who use three adjustable constants to represent the density variation of c.

To express V/n as a function of p and T we may solve equation 164 for V/n to any number of terms desired (16, 161). For the purpose of determining the variation of many of the thermodynamic properties of pure gases with pressure it is sufficient to use the expression:

$$\frac{V}{n} = \frac{RT}{p} + \frac{\beta}{RT} \tag{165}$$

$$\frac{\beta}{RT} = B_0 - \frac{A_0}{RT} - \frac{c}{T^3}$$

TABLE 1

Values of the constants of the Beattie–Bridgeman equation of state for several gases

$$p = [RT(1 - \epsilon)/V^2][V + B] - A/V^2$$

$$A = A_0(1 - a/V) \qquad B = B_0(1 - b/V) \qquad \epsilon = c/VT^3$$

Units: normal atmospheres; liters per mole; °K. (T°K. = t°C. + 273.13); R = 0.08206

GAS	A_0	a	B_0	b	$10^{-4} \times c$	MOLECULAR WEIGHT	REFERENCES
He	0.0216	0.05984	0.01400	0.0	0.0040	4.00	(19, 21)
Ne	0.2125	0.02196	0.02060	0.0	0.101	20.2	(19, 21)
A	1.2907	0.02328	0.03931	0.0	5.99	39.91	(19, 21)
H_2	0.1975	−0.00506	0.02096	−0.04359	0.0504	2.0154	(19, 21)
N_2	1.3445	+0.02617	0.05046	−0.00691	4.20	28.016	(19, 21)
O_2	1.4911	0.02562	0.04624	+0.004208	4.80	32	(19, 21)
Air	1.3012	0.01931	0.04611	−0.01101	4.34	28.964	(19, 21)
I_2	17.0		0.325		4000	253.864	(60)
CO_2	5.0065	0.07132	0.10476	+0.07235	66.00	44.000	(19, 21)
NH_3	2.3930	0.17031	0.03415	+0.19112	476.87	17.0311	(21, 27)
CH_4	2.2769	0.01855	0.05587	−0.01587	12.83	16.0308	(19, 21)
C_2H_4	6.1520	0.04964	0.12156	+0.03597	22.68	28.0308	(70, 21)
C_2H_6	5.8800	0.05861	0.09400	+0.01915	90.00	30.0462	(23)
C_3H_8	11.9200	0.07321	0.18100	+0.04293	120.00	44.0616	(26)
1-C_4H_8	16.6979	0.11988	0.24046	+0.10690	300.00	56.0616	(29)
iso-C_4H_8	16.9600	0.10860	0.24200	+0.08750	250.00	56.0616	(25)
n-C_4H_{10}	17.7940	0.12161	0.24620	+0.09423	350.00	58.077	(31)
iso-C_4H_{10}	16.6037	0.11171	0.23540	+0.07697	300.00	58.077	(30)
n-C_5H_{12}	28.2600	0.15099	0.39400	+0.13960	400.00	72.0924	(28)
neo-C_5H_{12}	23.3300	0.15174	0.33560	+0.13358	400.00	72.0924	(22)
n-C_7H_{16}	54.520	0.20066	0.70816	+0.19179	400.00	100.1232	(172)
CH_3OH	33.309	0.09246	0.60362	+0.09929	32.03	32.0308	(111, 114)
$(C_2H_5)_2O$	31.278	0.12426	0.45446	+0.11954	33.33	74.077	(19, 21)

GAS	$A_0^{1/2}$	$B_0^{1/3}$	$10^{-2} \times c^{1/2}$	GAS	$A_0^{1/2}$	$B_0^{1/3}$	$10^{-2} \times c^{1/2}$
He	0.14697	0.241014	0.0632	C_2H_4	2.48032	0.495371	4.7624
Ne	0.46098	0.274129	0.3178	C_2H_6	2.42487	0.454684	9.4868
A	1.13609	0.340017	2.4474	C_3H_8	3.45254	0.565665	10.9545
H_2	0.44441	0.275717	0.2245	1-C_4H_8	4.08631	0.621843	17.3205
N_2	1.15953	0.369529	2.0494	iso-C_4H_8	4.11825	0.623168	15.8114
O_2	1.22111	0.358927	2.1909	n-C_4H_{10}	4.21829	0.626752	18.7083
Air	1.14070	0.358590	2.0833	iso-C_4H_{10}	4.07476	0.617451	17.3205
I_2	4.12311	0.687534	63.2456	n-C_5H_{12}	5.31601	0.733104	20.0000
CO_2	2.23752	0.471410	8.1240	neo-C_5H_{12}	4.83011	0.694929	20.0000
NH_3	1.54693	0.324437	21.8374	n-C_7H_{16}	7.38377	0.891341	20.0000
CH_4	1.50894	0.382290	3.5819	CH_3OH	5.77139	0.845126	5.6595
				$(C_2H_5)_2O$	5.59267	0.768833	5.7732

Sartori (160) reports constants for water vapor for the temperature range 400–480° C.

This gives satisfactory results, except for calculation of the density itself, to moderate pressures (12, 13, 14, 60, 75, 76, 77, 79).

In order to compute values of the thermodynamic properties of a pure gas at various pressures we can proceed as follows:

1. For the most accurate values use equation 164 to evaluate the integrals of equations 97, 98, 106 through 110 and compute the pressure from equation 164. Note that equation 164 gives the value of pV which occurs in equations 106 and 108.
2. For more approximate results use equation 165 to evaluate the integrals of equations 88, 89, and 99 through 103. Now pV can be expressed by equation 165.

B. Relations among the equation of state constants and evaluation of the constants from critical data

From a study of the equation of state constants and critical constants of gases Keyes (93) found that the relation:

$$c = 0.023 \frac{A_0^3}{R^3 B_0^2}$$

held fairly well for non-polar gases. Hirschfelder and Roseveare (86) have interpreted the constants A_0, B_0, and c of equation 164 in terms of the energy of interaction and collision diameters of molecules by comparison of the second virial coefficient β/RT with that given by the 6–12 potential of the Lennard-Jones (115, 116) theory. They find 0.024 for the theoretical value of the coefficient. Corner (44) used a 6–9 potential and found the value 0.013. He extended the method to the third virial coefficient γ/RT which gave the theoretical relations:

$$a = 0.45 B_0; \qquad b = -0.1 B_0$$

The equation of state constants of a good many of the non-polar gases are in fair agreement with these relations. Stockmayer (174, 175) used an exponential in place of an algebraic temperature function in β/RT for polar gases and found a similar relation among B_0, A_0, and D (the constant replacing c).

The law of corresponding states originally arose through the prediction of a critical point by van der Waals' equation and the possibility of writing a reduced equation of state containing only universal constants. It fell into disrepute but recently interest has revived in both the practical and the theoretical aspects of the law (32). From equation 164 we note that the following are dimensionless ratios:

$$A_{0r} = \frac{A_0}{R^2 T_c^2/p_c}, \qquad B_{0r} = \frac{B_0}{RT_c/p_c}, \qquad c_r = \frac{c}{RT_c^4/p_c} \tag{166}$$

$$a_r = \frac{a}{RT_c/p_c}, \qquad b_r = \frac{b}{RT_c/p_c}$$

In these p_c and T_c are the critical pressure and temperature of the substance to which the constants apply. The relations are written in this form since the value of the critical volume, v_c, is never known so exactly as those of p_c and T_c, and the critical ratio $p_c v_c/RT_c$ is known to be approximately constant.

Keyes (93) found that the first three of these ratios for the substances whose equation of state constants were known fell into two groups, each surprisingly constant: one group of the non-polar gases and the other of the polar gases. Maron and Turnbull (136) and Brown (42) proposed to determine the equation of state constants of gases whose critical pressure and temperature are known but compressibility unknown by determining the values of A_{0r}, B_{0r}, c_r, a_r, b_r, for nitrogen and then use these to compute the parameters for other gases by equation 166. Although they obtained good results it might be preferable to consider the ratios somewhat variable and to evaluate them for a gas as nearly similar as possible to the gas for which the parameters are desired.

Su and Chang (176) made use of the relations of equation 166 and equation 164 to write the reduced equation:

$$p_r = \frac{T_r}{v_r} + \frac{\beta_r}{v_r^2} + \frac{\gamma_r}{v_r^3} + \frac{\delta_r}{v_r^4} \tag{167}$$

$$\beta_r = T_r B_{0r} - A_{0r} - c_r/T_r^2$$

$$\gamma_r = - T_r B_{0r} b_r + A_{0r} a_r - B_{0r} c_r/T_r^2$$

$$\delta_r = B_{0r} b_r c_r/T_r^2$$

$$p_r = p/p_c, \qquad v_r = vp_c/RT_c, \qquad T_r = T/T_c$$

They re-smoothed the compressibility data on seven hydrocarbons (methane, ethane, ethylene, propane, *n*-butane, isopentane, and *n*-heptane) to give p_r in terms of v_r and T_r, averaged p_r for the seven gases, and determined the values of A_{0r}, B_{0r}, c_r, a_r, and b_r in the usual manner. The average deviation of the calculated from the observed value of p_r was 0.7 per cent; and of the computed pressures from the observed pressures for the seven gases, 1 per cent. Application of this equation with the same value of the reduced constants to nine other gases (argon, helium, neon, hydrogen, oxygen, nitrogen, carbon dioxide, air, and ethyl ether) gave a representation of the pressures to 0.63 per cent.

In general, when the equation of state constants of a gas are not known but the critical pressure and temperature are known, we can use equation 167; or it is perhaps better to employ equation 164 or 165 with the values of the constants determined from equations 166 and the following values of Su and Chang:

$$A_{0r} = 0.4758 \qquad a_r = 0.1127 \tag{168}$$

$$B_{0r} = 0.18764 \qquad b_r = 0.03833$$

$$c_r = 0.05$$

Thus in equations 164 and 165:

$$A_0 = 0.4758 \left(\frac{RT_c}{p_c}\right)^2 p_c \qquad a = 0.1127 \left(\frac{RT^c}{p_c}\right) \tag{169}$$

$$B_0 = 0.18764 \left(\frac{RT_c}{p_c}\right) \qquad b = 0.03833 \left(\frac{RT_c}{p_c}\right)$$

$$= 0.05 \left(\frac{RT_c}{p_c}\right) T_c^3$$

C. The equation of state for gas mixtures

An equation of state for gas mixtures is an equation expressing the pressure of the mixture in terms of the volume, temperature, and the mole numbers and equation of state constants of the pure constituent gases. The earlier investigators felt that the equation of state of a gas mixture of constant composition had the same form as that of a pure gas and that each parameter of the mixture, say B_m, was related to the corresponding parameters of the pure gases by the relation:

$$B_m = \Sigma_i \Sigma_j x_i x_j B_{ij} \tag{170}$$

where x_i, x_j are mole fractions and the sums go over all of the species present in the mixture. The method of computing the interaction parameter B_{ij} for the interaction of unlike molecules from B_{ii} and B_{jj} has been the subject of a number of theoretical investigations (39, 59, 132, 181). The early workers in this field proposed the relations:

$$B_{ij} = \tfrac{1}{2}(B_i + B_j) \qquad \text{(linear combination)} \tag{171}$$

$$B_{ij} = B_i^{1/2} B_j^{1/2} \qquad \text{(square-root combination)} \tag{172}$$

$$B_{ij} = \tfrac{1}{8}(B_i^{1/3} + B_j^{1/3})^3 \qquad \text{(Lorentz combination)} \tag{173}$$

where B_i and B_j are written for B_{ii} and B_{jj}, respectively. The first relation gives the largest value of B_{ij} and the second, the smallest. Equation 172 was used for the a constant of van der Waals' equation of state and theory indicates that equation 173, averaging of molecular diameters, applies to the b constant, although equation 171, which gives nearly the same value when B_i and B_j do not differ greatly, has been frequently used for b.

The use of these combination rules in the van der Waals equation gave a moderately successful representation of the properties of gas mixtures (58, 101, 102, 103, 104, 105, 106, 107, 178, 179). The Keyes equation for gas mixtures gave better results (91, 92, 94, 95, 134), and that of Benedict, Webb, and Rubin, excellent results (35, 36, 38).

The use of combination of constants in equations 164 and 165 gives a good representation of the compressibilities of gas mixtures (11, 24, 34, 88), the expansion of gases on mixing (68, 69), and equilibria in gaseous phases (60, 73, 74, 75, 76, 77, 112, 113, 114).

Attempts have been made to write an equation of state for gas mixtures by finding combination rules for the virial coefficients (45, 50, 81, 82, 96, 144) or other properties of the mixture not involving the individual equation of state constants (5, 80, 98, 122). Su, Huang, and Chang (177) used the reduced equation 167 with the constants given in equations 168 and the pseudo-critical pressure and temperature of the mixture introduced by Kay (87) to calculate the pressures of gaseous mixtures.

The use of linear and quadratic combination rules for certain of the parameters of the Lennard-Jones (115, 116) expression for the second virial coefficient was found to give good representation of the properties of gas mixtures by Lennard-Jones and Cook (117), Hirschfelder and Roseveare (86), and Beattie and Stockmayer (33). The success of these investigations suggests that the second virial interaction coefficient β_{12}/RT be written in terms of B_{012}, A_{012}, and c_{12} of equation 164 as:

$$\frac{\beta_{12}}{RT} = B_{012}\left[1 - \frac{\phantom{A_{012}}}{B_{012}RT} \quad \frac{c_{12}}{B_{012}T^3}\right] \tag{174}$$

and that the B_{012} outside the brackets be obtained by Lorentz combination, and (A_{012}/B_{012}) and (c_{12}/B_{012}) by square-root combination.

Stockmayer (175) showed that A_0 should be split into two terms for a polar gas: one corresponding to the attractive constant A_0 for a non-polar gas and the other depending on the dipole moment. Each is to be combined separately by square-root combination for a gas mixture containing polar gases.

Practical considerations require that an equation of state for gas mixtures involve fairly simple combination rules for the constants. A study (34) of various methods of combining the constants A_0, B_0, and c of equation 164 for the system methane–n-butane indicated that the best representation of the compressibilities of the mixtures are obtained by square-root combination for A_{012} and c_{12}, and Lorentz combination for B_{012}. The average percentage deviation of the computed from the observed pressures was 0.74 per cent, about half of which is caused by the failure of equation 164 to represent the pressures of the pure gases. The simpler rules, square-root for A_{012} and linear for B_{012} and c_{12}, represented the measurements with about twice as large an average percentage deviation. In this study linear combination was used for a_{12} and b_{12}.

Using square-root combination (equation 172) for A_{012} and c_{12}, Lorentz combination (equation 173) for B_{012}, and linear combination (equation 171) for a_{12} and b_{12} in equation 170 to obtain the parameters for a gas mixture, we find from equation 164 the following equations of state for $\Sigma_i n_i$ moles of a gas mixture:

$$p = \frac{\Sigma_i n_i RT}{V} + \frac{(\Sigma_i n_i)^2 \beta_m}{V^2} + \frac{(\Sigma_i n_i)^3 \gamma_m}{V^3} + \frac{(\Sigma_i n_i)^4 \delta_m}{V^4} \tag{175}$$

where

$$\begin{aligned} \beta_m &= RTB_{0m} - A_{0m} - \frac{Rc_m}{T^2} \\ \gamma_m &= -RTB_{0m}b_m + A_{0m}a_m - \frac{RB_{0m}c_n}{T^2} \\ \delta_m &= \frac{RB_{0m}b_m c_n}{T^2} \end{aligned} \tag{176}$$

and

$$\left.\begin{aligned} A_{0m} &= (\Sigma_i x_i A_{0i}^{1/2})^2 \\ B_{0m} &= \tfrac{1}{8}\Sigma_i\Sigma_j x_i x_j (B_{0i}^{1/3} + B_{0j}^{1/3})^3 = \tfrac{1}{4}(\Sigma_i x_i B_{0i}) + \tfrac{3}{4}(\Sigma_i x_i B_{0i}^{1/3})(\Sigma_i x_i B_{0i}^{2/3}) \\ c_m &= (\Sigma_i x_i c_i^{1/2})^2 \\ a_m &= \Sigma_i x_i a_i \\ b_m &= \Sigma_i x_i b_i \end{aligned}\right\} \tag{177}$$

The approximate equation 165 becomes for a gas mixture:

$$V = \frac{\Sigma_i n_i RT}{p} + \frac{(\Sigma_i n_i)\beta_m}{RT} \tag{178}$$

where β_m is given by equations 176 and 177.

Equations 175 and 178 in conjunction with equations 176 and 177 purport to represent the pressure [or volume] of a gas mixture as a function of V, T, n_1, n_2, $\cdots$ [or p, T, n_1, n_2, $\cdots$] containing parameters of the pure gases 1, 2, $\cdots$ only. No measurements on mixtures are necessary. These relations can be used to evaluate the integrals of Section XIII, paragraphs E, F, and G.

The general limit method gives a partial though exact solution of the problem of expressing the thermodynamic properties of real gases and gas mixtures in terms of those of pure gases. The general limit method and combination of constants in an equation of state give a complete though approximate solution to the problem. The approximation is introduced through the equation of state and the combination rules.

D. Integrated approximate equations for pure gases

The definite integrals of the general relations of Section XII F (page 166) can be expressed in terms of the equation of state constants of a gas by evaluation of the integrands from equation 164. The resulting equations state the thermodynamic properties of gases in terms of the independent variables V and T. Less exact but simpler expressions in terms of p and T may be derived from the use of equation 165 in the relations of Section XII E (page 165). These are sufficient for many purposes and are given below.

$$E = -\left[\frac{A_0}{RT} + \frac{3c}{T^3}\right] np - nRT + nh^0 \tag{179}$$

$$H = \left[B_0 - \frac{2A_0}{RT} - \frac{4c}{T^3}\right] np + nh^0 \tag{180}$$

$$S = -nR \ln p - \left[\frac{A_0}{RT^2} + \frac{3c}{T^4}\right] np + ns^0 \tag{181}$$

$$F_{VT} = nRT \ln p - nRT + nh^0 - nTs^0 \tag{182}$$

$$F_{pT} = nRT \ln p + \left[B_0 - \frac{A_0}{RT} - \frac{c}{T^3}\right] np + nh^0 - nTs^0 \tag{183}$$

$$\mu = RT \ln p + \left[B_0 - \frac{A_0}{RT} - \frac{c}{T^3}\right] p + h^0 - Ts^0 \tag{184}$$

$$RT \ln f = RT \ln p + \left[B_0 - \frac{A_0}{RT} - \frac{c}{T^3}\right] p \tag{185}$$

From the usual thermodynamic relations and equation 165 we find for the heat capacities and the Joule and Joule–Thompson coefficient the following relations:

$$C_V = nc_V^* + \frac{6c}{T^4}\,np \tag{186}$$

$$\left[\frac{2A_0}{RT^2} + \frac{12c}{T^4}\right]. \tag{187}$$

$$C_p - C_V = nR + \left[\frac{2A_0}{RT^2} + \frac{6c}{T^4}\right]np \tag{188}$$

$$C_V\left(\frac{\partial T}{\partial V}\right)_E = -\left(\frac{\partial E}{\partial V}\right)_T = \left[\frac{\partial(p/T)}{\partial(1/T)}\right]_V \tag{189}$$

$$= -\left[\frac{A_0}{R^2T^2} + \frac{3c}{RT^4}\right]p^2 \tag{190}$$

$$C_p\left(\frac{\partial T}{\partial p}\right)_H = -\left(\frac{\partial H}{\partial p}\right)_T = -\left[\frac{\partial(V/T)}{\partial(1/T)}\right]_p \tag{191}$$

$$= n\left[-B_0 + \frac{2A_0}{RT} + \frac{4c}{T^3}\right] \tag{192}$$

In the absence of compressibility data from which A_0, B_0, and c can be determined, we can estimate the values of these constants for use in the above equations from the critical pressure and temperature by means of equations 169.

The above relations do not include any of the terms involving the third or higher virial coefficients. The expansion of equation 164 ending at the term in p^2:

$$\frac{V}{n} = \frac{RT}{p} + \frac{\beta}{RT} + \left[\frac{\gamma}{(RT)^2} \qquad \frac{}{(RT)^3}\right]p$$
$$+ \left[\frac{\delta}{(RT)^3} - \frac{3\beta\gamma}{(RT)^4} + \frac{2\beta^3}{(RT)^5}\right]p^2 \tag{193}$$

proposed by Scatchard (161) reproduces the density to moderately high pressure and may be employed to extend equations 179–192 to higher pressures.

E. Integrated approximate equations for gas mixtures

The most accurate results are obtained from the general relations of Section XIII F (page 171) and the complete equation of state for gas mixtures, equation 175. These express the properties of gas mixtures in terms of the independent variables V, T, and the mole numbers. The equations given below come from the approximate equation (equation 178) and the general relations of Section XIII E; they express the properties of gas mixtures in terms of p, T, and the mole numbers.

If we write the second virial coefficient β_m of a gas mixture in the form of equation 170, β_{ii} and β_{jj} are the virial coefficients of the pure gases i and j, and the interaction coefficient β_{ij} is by equations 176 and 177:

$$\beta_{ij} = \frac{RT}{8}\left(B_{0i}^{1/3} + B_{0j}^{1/3}\right)^3 - A_{0i}^{1/2}A_{0j}^{1/2} - \frac{Rc_i^{1/2}c_j^{1/2}}{T^2} \tag{194}$$

Let

$$D_{ij} = 2\beta_{ij} - \beta_{ii} - \beta_{jj} \tag{195}$$

then

$$\frac{D_{ij}}{RT} = -\left[\frac{3}{4}\left(B_{0i}^{1/3} - B_{0j}^{1/3}\right)\left(B_{0i}^{2/3} - B_{0j}^{2/3}\right) - \frac{\left(A_{0i}^{1/2} - A_{0j}^{1/2}\right)^2}{RT} - \frac{\left(c_i^{1/2} - c_j^{1/2}\right)^2}{T^3}\right] \tag{196}$$

where

$$D_{ii} = D_{jj} = 0, \qquad D_{ij} = D_{ji} \tag{197}$$

For each pair of gases D_{ij} is an algebraic function of temperature and the coefficients are computed from the A_0, B_0, and c constants for the two gases. At constant temperature D_{ij} is simply a number for a given pair of gases.

The expression:

$$\left(\frac{\partial V}{\partial n_i}\right)_{pTn} - \frac{RT}{p} = \frac{1}{RT}\left[\frac{\partial(\Sigma_i n_i)\beta_m}{\partial n_i}\right]_{Tn}$$

$$= \frac{1}{RT}\left[2\Sigma_j x_j \beta_{ij} - \Sigma_i\Sigma_j x_i x_j \beta_{ij}\right] \tag{198}$$

$$= \frac{1}{RT}\left[\beta_{ii} + \Sigma_j x_j D_{ij} - \tfrac{1}{2}\Sigma_i\Sigma_j x_i x_j D_{ij}\right] \tag{199}$$

$$= \frac{1}{RT}\left(\beta_i + D_i\right) \tag{200}$$

where the summations extend over all species present in the gas mixture, β_i has been written for the second virial coefficient β_{ii} of the pure gas i, and D_i/RT is the difference between the partial molal volume of gas i in the mixture and its molal volume at the same pressure and temperature as computed by equation 165:

$$\frac{D_i}{RT} = \left(\frac{\partial V}{\partial n_i}\right)_{pTn} - \left(\frac{RT}{p} + \frac{\beta_i}{RT}\right) \tag{201}$$

In general D_i is much smaller than β_i, and it depends on the temperature, composition, and equation of state constants of all gases present in the mixture. We may write D_i in the following forms:

$$\frac{D_i}{RT} = \frac{1}{RT}\left(\Sigma_j x_j D_{ij} - \tfrac{1}{2}\Sigma_i \Sigma_j x_i x_j D_{ij}\right) \tag{202}$$

$$= -\left[\frac{3}{4}\left(B_{0i}^{1/3} - \Sigma_i x_i B_{03}^{1/3}\right)\left(B_{0i}^{2/3} - \Sigma_i x_i B_{0i}^{2/3}\right) - \frac{\left(A_{0i}^{1/2} - \Sigma_i x_i A_{0i}^{1/2}\right)^2}{RT} - \frac{\left(c_i^{1/2} - \Sigma_i x_i c_i^{1/2}\right)^2}{T^3}\right] \tag{203}$$

where the summations extend over all gases present.

For a gas 1 in a binary mixture:

$$D_1 = D_{12}x_2^2 \tag{204}$$

and for a gas 1 in a ternary mixture:

$$D_1 = D_{12}x_2^2 + D_{13}x_3^2 + (D_{12} + D_{13} - D_{23})x_2x_3 \tag{205}$$

The thermodynamic properties of gas mixtures follow.

$$E = -\left[\frac{A_{0m}}{RT} + \frac{3c_m}{T^3}\right](\Sigma_i n_i)p - (\Sigma_i n_i)RT + (\Sigma_i n_i h_i^0) \tag{206}$$

$$H = \left[B_{0m} - \frac{2A_{0m}}{RT} - \frac{4c_m}{T^3}\right](\Sigma_i n_i)p + (\Sigma_i n_i h_i^0) \tag{207}$$

$$S = -(\Sigma_i n_i R \ln px_i) - \left[\frac{A_{0m}}{RT^2} + \frac{3c_m}{T^4}\right](\Sigma_i n_i)p + (\Sigma_i n_i s_i^0) \tag{208}$$

$$F_{VT} = (\Sigma_i n_i RT \ln px_i) - (\Sigma_i n_i)RT + (\Sigma_i n_i h_i^0) - T(\Sigma_i n_i s_i^0) \tag{209}$$

$$F_{pT} = (\Sigma_i n_i RT \ln px_i) + \left[B_{0m} - \frac{A_{0m}}{RT} - \frac{c_m}{T^3}\right](\Sigma_i n_i)p + (\Sigma_i n_i h_i^0) - T(\Sigma_i n_i s_i^0) \tag{210}$$

$$\mu_i = RT \ln px_i + \left(\frac{\beta_i}{RT} + \frac{D_i}{RT}\right)p + h_i^0 - Ts_i^0 \tag{211}$$

$$RT \ln f_i = RT \ln px_i + \left(\frac{\beta_i}{RT} + \frac{D_i}{RT}\right)p \tag{212}$$

$$C_V = (\Sigma_i n_i c_{vi}^*) + \frac{6c_m}{T^4}(\Sigma_i n_i)p \tag{213}$$

$$C_p = (\Sigma_i n_i c_{pi}^*) + \left[\frac{2A_{0m}}{RT^2} + \frac{6c_m}{T^4}\right](\Sigma_i n_i)p \tag{214}$$

$$C_p - C_V = (\Sigma n_i)R + \left[\frac{2A_{0m}}{RT^2} + \frac{6c_m}{T^4}\right](\Sigma_i n_i)p \tag{215}$$

For the Joule and Joule–Thompson coefficients we obtain the same relations as equations 190 and 192 with the constants of the pure gases replaced by those of the mixture and n replaced by $\Sigma_i n_i$.

F. Integrated approximate equilibrium relationships

From the general equilibrium relations of Section XIII G (page 171) we find the following equations:

1. *Mass action functions for homogeneous gaseous reactions:* Equations 151 and 153 become:

$$\ln\frac{K_p}{K_p^*} = -\left[\frac{\Sigma_i \nu_i \beta_i}{RT} + \frac{\Sigma_i \nu_i D_i}{RT}\right]\frac{p}{RT} \tag{216}$$

$$= -\left\{\Sigma_i \nu_i B_{0i} - \frac{\Sigma_i \nu_i A_{0i}}{RT} - \frac{\Sigma_i \nu_i c_i}{T^3}\right\} \frac{p}{RT}$$

$$+ \left\{ \tfrac{3}{4}[\Sigma_i \nu_i (B_{0i}^{1/3} - \Sigma_i x_i B_{0i}^{1/3})(B_{0i}^{2/3} - \Sigma_i x_i B_{0i}^{2/3})] - \frac{[\Sigma_i \nu_i (A_{0i}^{1/2} - \Sigma_i x_i A_{0i}^{1/2})^2]}{RT} - \frac{[\Sigma_i \nu_i (c_i^{1/2} - \Sigma_i x_i c_i^{1/2})^2]}{T^3}\right\} \frac{p}{RT} \tag{217}$$

$$\ln K_p^* = \frac{1}{R}\left[\int_{T_0}^{T} \frac{(\Sigma_i \nu_i c_{pi}^*)}{T}\, \mathrm{d}T\right] - \frac{1}{RT}\left[\int_{T_0}^{T} (\Sigma_i \nu_i c_{pi}^*)\, \mathrm{d}T\right] - \frac{1}{RT}(\Sigma_i \nu_i h_{0i}^0) + \frac{1}{R}(\Sigma_i \nu_i s_{0i}^0) \tag{218}$$

$$= \int_{T_0}^{T}\left[\int_{T_0}^{T} (\Sigma_i \nu_i c_{pi}^*)\, \mathrm{d}T\right]\frac{\mathrm{d}T}{T^2} - \frac{1}{RT}(\Sigma_i \nu_i h_{0i}^0) + \frac{1}{R}(\Sigma_i \nu_i s_{0i}^0) \tag{219}$$

All summations extend over all species in the mixture, the stoichiometric coefficient ν being zero for a gas that does not enter into the chemical reaction.

Equations 217 and 219 express K_p in terms of the temperature, pressure, composition of the equilibrium mixture, the equation of state constants for all gases present, the heat capacities of the gases entering into the reaction, and the heat content and entropy constants of the gases entering into the reaction. The last two sums of equation 219 may be considered adjustable in the absence of heat content and entropy data on the gases. Although equation 217 was derived for moderate pressures, it applies through a compensation of errors to gaseous equilibria at high pressure. Gillespie and Beattie (75, 76, 77) found it possible to correlate the measurements of the ammonia synthesis equilibrium to 1000 atm. including equilibrium mixtures containing argon and those containing no argon. They used linear combination for B_{012} and c_{12}.

To compute the composition of the equilibrium mixture given the initial compositions, the pressure, the temperature, and the necessary constants for the pure gases we first calculate K_p^*. From K_p^* (the ideal gas value of K_p) compute the composition of the equilibrium mixture. Use these values of the mole fractions in the right-hand side of equation 217 to calculate a provisional value of K_p from which provisional values of the compositions can be obtained. Repeat until no change in the calculated compositions results. In general two approximations are sufficient even at high pressures.

2. *Equilibrium pressure of a gas in a mixture:* Equation 156 gives:

$$\ln \frac{p_{ek}}{px_i} = \frac{\beta_i}{RT}\left(\frac{p - p_{ek}}{RT}\right) + \frac{D_i}{RT}\frac{p}{RT} \qquad i \equiv k = 1, 2, \cdots \tag{220}$$

If we express D_i in terms of D_{ij}, we find for a gas 1 in a ternary mixture:

$$\ln \frac{p_{e1}}{px_1} = \frac{\beta_1}{RT}\left(\frac{p - p_{e1}}{RT}\right) + \frac{1}{RT}[D_{12}x_2^2 + D_{13}x_3^2 + (D_{12} + D_{13} - D_{23})x_1x_2]\frac{p}{RT} \tag{221}$$

For a binary mixture all terms in the bracket except the first disappear. The general equation 156 was derived (10) a number of years ago, and the approximate equation 220 was obtained by Gillespie (69).

To compute the equilibrium pressure of a gas given the total pressure and composition of the mixture, place p_{ek} on the right-hand side of the equation equal to its ideal gas value px_i for the first approximation. Then repeat the calculation with the provisional value of p_{ek} resulting from the first approximation.

To compute the composition of a mixture of n gases given the total pressure and $n - 1$ values of p_{ek}, use the ideal gas values $x_i = p_{ek}/p$ for the compositions occurring on the right-hand side of the equation. Repeat if necessary.

3. *Effect of an inert gas on a univariant equilibrium:* Let

$$\bar{v}_{c1}(p,T) = \bar{v}_{c1}^*(0,T)(1 - \kappa_1 p) \tag{222}$$

where $\bar{v}_{c1}^*$ is the (extrapolated) value of $\bar{v}_{c1}$ at zero pressure and κ_1 is the compressibility of this quantity.

The Poynting equation (equation 160) gives:

$$\ln \frac{p_{e1}}{p_{01}} = -\frac{\beta_1}{RT}\left(\frac{p_{e1} - p_{01}}{RT}\right) + \bar{v}_{c1}^*\left(\frac{p - p_{01}}{RT}\right) - \frac{\kappa_1}{2}\left(\frac{p^2 - p_{0i}^2}{RT}\right) \tag{223}$$

where p_{01} is the equilibrium pressure of the condensed complex under its own vapor pressure and p_{e1} is the equilibrium pressure under the pressure p. When the pressure p is obtained by forcing an inert gas insoluble in the condensed complex into the system, we find from equation 161:

$$\ln \frac{px_1}{p_{01}} = \left(\bar{v}_{c1}^* - \frac{\beta_1}{RT}\right)\left(\frac{p - p_{01}}{RT}\right) - \frac{D_1}{RT}\frac{p}{RT} - \frac{\kappa_1}{2}\left(\frac{p^2 - p_{01}^2}{RT}\right) \tag{224}$$

$$= \left(\bar{v}_{c1}^* - \frac{\beta_1}{RT}\right)\left(\frac{p - p_{01}}{RT}\right) - \frac{D_{12}x_2^2}{RT}\frac{p}{RT} - \frac{\kappa_1}{2}\left(\frac{p^2 - p_{01}^2}{RT}\right) \tag{225}$$

In equation 225 x_2 is the mole fraction of the inert gas in the gas phase. If the inert gas is soluble in the condensed complex, equations 226 and 227 should be used.

Equation 224 combines equation 220 for the equilibrium pressure of a gas in a mixture and equation 223 for the Poynting effect. It was derived by Gerry and Gillespie (60).

Equation 225 is used to compute p_{01} given p and x_1, or to compute x_1 given p and p_{01} in the manner described under equation 221.

4. *Equilibrium between gaseous mixtures and condensed solutions:* Equation 163 gives:

$$\ln \frac{px_i}{p_{0i}y_i} = \left(\bar{v}_{ci}^* - \frac{\beta_i}{RT}\right)\left(\frac{p - p_{0i}}{RT}\right) - \frac{D_i}{RT}\frac{p}{RT} - \frac{\kappa_i}{2}\left(\frac{p^2 - p_{0i}^2}{RT}\right) + \ln \gamma_i \tag{226}$$

For component 1 of a binary system:

$$\ln \frac{px_1}{p_{01}y_1} = \left(\bar{v}_{c1}^* - \frac{\beta_1}{RT}\right)\left(\frac{p - p_{01}}{RT}\right) - \frac{D_{12}x_2^2}{RT}\frac{p}{RT} - \frac{\kappa_1}{RT}\left(\frac{p^2 - p_{01}^2}{RT}\right) + \ln \gamma_1 \tag{227}$$

This equation was derived by Scatchard and Raymond (165). For the use of equation 227 see under equation 225.

The relations given in the last three paragraphs have been applied to numerous equilibria. Lurie and Gillespie (134) used an equation similar to equation 224 to correlate their measurements on the equilibrium $BaCl_2 \cdot 8NH_3(s)$, $BaCl_2(s)$, $NH_3(g) + N_2(g)$ and applied it to the data of Pollitzer and Strebel (145) on the system $CO_2(l)$, $CO_2(g) + N_2(g)$; they used combination of constants in the Keyes equation of state. Gerry and Gillespie (60) applied the relation to the correlation of the data on the vapor pressure of iodine determined by the gas stream method. Apparent discrepancies caused by the use of different inert gases were resolved. Scatchard and coworkers (164–170) and Benedict and coworkers (35) applied equation 227 to liquid–vapor equilibria at about 1 atm. pressure. Here the pressures were low and the parameters D_{12} and κ_1 could be placed equal to zero; the main effect was that introduced by γ_1. Goff and coworkers (81, 82) applied an equation similar to equation 225 to moist air; they used a combination of second virial coefficients. Stockmayer (175) applied equation 226 to the data of Saddington and Krase (150) and Bartlett (3) on the system $H_2O(l)$, $H_2O(g) + N_2(g)$ and to unpublished data on the system $H_2O(l)$, $H_2O(g) + CO_2(g)$; he used a special method for combining A and D (corresponding to A_0 and roughly to Rc/A_0) for the polar gases.

The effects predicted by the ideal gas law may be obtained by placing the parameters β and D equal to zero in equations 220, 223, 224, and 226. In general the mole fraction of a vapor in a gas mixture is greater than that computed by Dalton's law (sometimes as much as 100 times), but it may be less if the inert gas is above its Boyle temperature.

G. The energy and entropy constants

From equations 88 and 104 we find that the heat content constant h_0^0 is the molal heat content of a gas at T_0 and a very low pressure p^*; from equations 97 and 111, e_0^0 is the molal energy at T_0 and p^*; and from equations 89 and 105 $s_0^0 - R \ln p^*$ is the molal entropy at T_0 and p^*. The parameters h^0, e^0, and $s^0 - R \ln p^*$ are the corresponding molal quantities at T and p^*. If we choose a standard state of 1 atm. and T_0, we note that h_0^0, e_0^0, and s_0^0 are approximately equal to the molal heat content, energy, and entropy of the gas in the standard state (see equations 179, 180, and 181).

The evaluation of h_0^0 and s_0^0 ($e_0^0 = h_0^0 - RT_0$) depends on what changes in state of the gas mixture are to be studied, and may be made on the basis of the following considerations:

1. *Chemical reactions may occur among the constituent gases:* Arbitrary values, say zero, may be assigned to the molal heat content and entropy of all elementary substances in the standard state p_0, T_0. The constants h_0^0 and s_0^0 for such elements as are gases under these conditions can be computed from equations 180 and 181, and for elements that are not gases and gases that are not elements can be calculated from these equations and other thermodynamic data. Thus from the usual relations we can compute the heat content and entropy of an

element in the condensed state at its vapor pressure and T_0, add the heat content and entropy of vaporization to give these quantities for the vapor at the vapor pressure and T_0, and apply equations 180 and 181 to determine h_0^0 and s_0^0. For a gaseous compound we can compute $\Sigma_i\nu_i h^0_{0i}$ and $\Sigma_i\nu_i s^0_{0i}$ from equations 217 and 219 from equilibrium measurements of a reaction in which the compound is formed from its elements or of some reaction involving the compound with other substances whose constants are known. This has been done for the ammonia synthesis reaction (75, 76).

When interested only in the effect of pressure on chemical equilibrium at constant temperature, we can use equation 217 with K_p^* considered to be an adjustable parameter.

2. *No chemical reaction occurs:* We can assign arbitrary values to the molal heat content and entropy of each individual substance or indeed to the mixture if the composition is constant. If the temperature is to vary the values are assigned only at p_0,T_0, but if the temperature is to be constant for all variations we can assign arbitrary values at p_0 and any temperature.

XV. COMPARISON OF THE GENERAL LIMIT METHOD WITH THE LEWIS AND RANDALL RULE AND THE GIBBS–DALTON LAW

The Lewis and Randall rule requires additivity of volumes at constant total pressure and temperature (equation 40). The Gibbs–Dalton law requires the additivity of pressures at constant total volume and temperature (equation 6). The general limit method places no such severe restrictions on the equation of state for gas mixtures; it does require that the pV product for a gas mixture approach $\Sigma_i n_i RT$ in a certain manner as the pressure approaches zero (equations 117).

The Lewis and Randall rule, the Gibbs–Dalton law, and the general limit method together with combination of constants in an equation of state each give a complete solution to the problem of computing the general thermodynamic properties of gas mixtures from data on the pure constituent gases alone. The latter gives the most accurate representation of the properties of gas mixtures; the Lewis and Randall rule and the Gibbs–Dalton law frequently give calculated values that bracket the measured value, one giving a high and the other a low value.

The Lewis and Randall rule requires that the partial molal volume of a gas in a mixture equal the molal volume of the pure gas measured at the temperature and total pressure of the mixture. It gives in place of equation 200 the relation:

$$\left(\frac{\partial V}{\partial n_i}\right)_{pTn} - \frac{RT}{p} = \frac{\beta_i}{RT} \tag{228}$$

that is, it requires that $D_i = 0$. If we delete all terms containing any mole fractions on the right-hand side of the equations of Section XIV F (page 183) we obtain the approximate equilibrium relationships given by the Lewis and Randall rule, the approximation being that of equation 165 for pure gases. Thus the rule predicts that K_p, p_{ei}/px_i, etc. do not depend on the composition of the gas mixture but only on the properties of pure gases.

The requirement of the Gibbs–Dalton law is given in equation 9. From equation 175 we find, on dropping terms after the second:

$$\left(\frac{\partial p}{\partial n_i}\right)_{VTn} - \frac{RT}{V} = (\beta_i + \beta_m + D_i)\frac{\Sigma_i n_i}{V^2} \qquad (i \equiv k = 1, 2, \cdots) \qquad (229)$$

To the same number of terms the Gibbs–Dalton law predicts:

$$\left(\frac{\partial p}{\partial n_i}\right)_{VTn} - \frac{RT}{p} = 2\beta_i \frac{n_{ck}}{V^2} \qquad (230)$$

where V and T have the same values in equations 229 and 230, and $n_{ck} = n_i$, the number of moles of gas i in $\Sigma_i n_i$ moles of the mixture.

Equation 229 used in equation 157 gives for the equilibrium concentration C_{ek} of a gas in a mixture:

$$\ln \frac{C_{ek}}{C_i} = (\beta_i + \beta_m + D_i)\frac{C}{RT} - 2\beta_i \frac{C_{ek}}{RT} \qquad (i \equiv k = 1, 2, \cdots) \qquad (231)$$

where C is the concentration of the mixture and C_i the concentration of gas i in the mixture. The Gibbs–Dalton law requires that the right-hand side of equation 231 be zero.

XVI. REFERENCES

(1) Amagat, E. H.: Ann. chim. phys. [5] **19**, 384 (1880).
(2) Amagat, E. H.: Compt. rend. **127**, 88 (1898).
(3) Bartlett, E. P.: J. Am. Chem. Soc. **49**, 65 (1927).
(4) Bartlett, E. P.: J. Am. Chem. Soc. **49**, 687, 1955 (1927).
(5) Bartlett, E. P., Cupples, H. L., and Tremearne, T. H.: J. Am. Chem. Soc. **50**, 1275 (1928).
(6) Baxter, G. P., and Grose, M. R.: J. Am. Chem. Soc. **37**, 1061 (1915).
(7) Baxter, G. P., Hickey, C. H., and Holmes, W. C.: J. Am. Chem. Soc. **29**, 127 (1907).
(8) Beattie, J. A. Phys. Rev. **31**, 680 (1928).
(9) Beattie, J. A. Phys. Rev. **32**, 691 (1928).
(10) Beattie, J. A. Phys. Rev. **32**, 699 (1928).
(11) Beattie, J. A. J. Am. Chem. Soc. **51**, 19 (1929).
(12) Beattie, J. A. Phys. Rev. **34**, 1615 (1929).
(13) Beattie, J. A. Phys. Rev. **35**, 643 (1930).
(14) Beattie, J. A. J. Math. Phys. **9**, 11 (1930).
(15) Beattie, J. A. Phys. Rev. **36**, 132 (1930).
(16) Beattie, J. A. Proc. Natl. Acad. Sci. U. S. **16**, 14 (1930).
(17) Beattie, J. A., and Bridgeman, O. C.: J. Am. Chem. Soc. **49**, 1665 (1927).
(18) Beattie, J. A., and Bridgeman, O. C.: J. Am. Chem. Soc. **50**, 3133 (1928).
(19) Beattie, J. A., and Bridgeman, O. C.: Proc. Am. Acad. Arts Sci. **63**, 229 (1928).
(20) Beattie, J. A., and Bridgeman, O. C.: J. Am. Chem. Soc. **50**, 3151 (1928).
(21) Beattie, J. A., and Bridgeman, O. C.: Z. Physik **62**, 95 (1930).
(22) Beattie, J. A., Douslin, D. R., and Levine, S. W.: To be published.
(23) Beattie, J. A., Hadlock, C., and Poffenberger, N.: J. Chem. Phys. **3**, 93 (1935).
(24) Beattie, J. A., and Ikehara, S.: Proc. Am. Acad. Arts Sci. **64**, 127 (1930).
(25) Beattie, J. A., Ingersoll, H. G., and Stockmayer, W. H.: J. Am. Chem. Soc. **64**, 548 (1942).
(26) Beattie, J. A., Kay, W. C., and Kaminsky, J.: J. Am. Chem. Soc. **59**, 1589 (1937).
(27) Beattie, J. A., and Lawrence, C. K.: J. Am. Chem. Soc. **52**, 6 (1930).

(28) Beattie, J. A., Levine, S. W., and Douslin, D. R.: To be published.
(29) Beattie, J. A., and Marple, S., Jr.: To be published.
(30) Beattie, J. A., Marple, S., Jr., and Edwards, D. G.: To be published.
(31) Beattie, J. A., Simard, G. L., and Su, G.-J.: J. Am. Chem. Soc. **61**, 26 (1939).
(32) Beattie, J. A., and Stockmayer, W. H.: Phys. Soc. (London), Reports on Progress in Physics **7**, 195 (1940).
(33) Beattie, J. A., and Stockmayer, W. H.: J. Chem. Phys. **10**, 473, 476 (1942).
(34) Beattie, J. A., Stockmayer, W. H., and Ingersoll, H. G.: J. Chem. Phys. **9**, 871 (1941).
(35) Benedict, M., Johnson, C. A., Solomon, E., and Rubin, L. C: Trans. Am. Inst. Chem. Engrs. **41**, 371 (1945).
(36) Benedict, M., Solomon, E., and Rubin, L. C.: Ind. Eng. Chem. **37**, 55 (1945).
(37) Benedict, M., Webb, G. B., and Rubin, L. C.: J. Chem. Phys. **8**, 334 (1940).
(38) Benedict, M., Webb, G. B., and Rubin, L. C.: J. Chem. Phys. **10**, 747 (1942).
(39) Berthelot, D.: Compt. rend. **126**, 1703, 1857 (1898).
(40) Braune, H., and Strassmann, F.: Z. physik. Chem. **A143**, 225 (1929).
(41) Bridgeman, O. C.: Phys. Rev. **34**, 527 (1929).
(42) Brown, F. W.: Ind. Eng. Chem. **33**, 1536 (1941).
(43) Brown, G. G.: Ind. Eng. Chem. **33**, 263 (1941).
(44) Corner, J.: Trans. Faraday Soc. **37**, 358 (1941).
(45) Corner, J.: Proc. Roy. Soc. (London) **57**, 737 (1946).
(46) Cupples, H. L.: J. Am. Chem. Soc. **51**, 1026 (1929).
(47) Dalton, J.: Nicholson's J. Sci. **5**, 241 (1801).
(48) Dalton, J.: Gilbert's Ann. Physik **12**, 385 (1802); **15**, 1 (1803).
(49) Deming, W. E., and Deming, L. S.: Phys. Rev. **56**, 108 (1939).
(50) Deming, W. E., and Shupe, L. E.: J. Franklin Inst. **208**, 389 (1929).
(51) Deming, W. E., and Shupe, L. E.: Phys. Rev. **37**, 638 (1931).
(52) Deming, W. E., and Shupe, L. E.: Phys. Rev. **38**, 2245 (1931).
(53) Deming, W. E., and Shupe, L. E.: Phys. Rev. **40**, 848 (1932).
(54) Deschner, W. W., and Brown, G. G.: Ind. Eng. Chem. **32**, 836 (1940).
(55) Dodge, B. F., and Newton, R. H.: Ind. Eng. Chem. **29**, 778 (1937).
(56) Donder, T. de: Compt. rend. **180**, 1922 (1925).
(57) Donder, T. de: *L'Affinité.* Gauthier-Villers, Paris (1927).
(58) Eucken, A., and Bresler, F.: Z. physik. Chem. **A134**, 130 (1928).
(59) Galitzine, B.: Ann. Physik **41**, 770 (1890).
(60) Gerry, H. T., and Gillespie, L. J.: Phys. Rev. **40**, 269 (1932).
(61) Gibbs, J. W.: *Scientific Papers*, Vol. I, pp. 155–8. Longmans, Green and Company, New York (1906).
(62) Reference 61, p. 165.
(63) Gibson, G. E., and Sosnick, B.: J. Am. Chem. Soc. **49**, 2172 (1927).
(64) Gillespie, L. J. J. Math. Phys. **4**, 84 (1925).
(65) Gillespie, L. J. Proc. Natl. Acad. Sci. U. S. **11**, 73 (1925).
(66) Gillespie, L. J. J. Am. Chem. Soc. **47**, 305 (1925)'
(67) Gillespie, L. J. J. Am. Chem. Soc. **48**, 28 (1926).
(68) Gillespie, L. J. Phys. Rev. **34**, 352 (1929).
(69) Gillespie, L. J. Phys. Rev. **34**, 1605 (1929).
(70) Gillespie, L. J. J. Phys. Chem. **33**, 354 (1929).
(71) Gillespie, L. J. Proc. Am. Acad. Arts Sci. **66**, 153 (1930).
(72) Gillespie, L. J. Phys. Rev. **36**, 121 (1930).
(73) Gillespie, L. J. Chem. Revs. **18**, 359 (1936).
(74) Gillespie, L. J.: Chem. Revs. **29**, 525 (1941).
(75) Gillespie, L. J., and Beattie, J. A.: Phys. Rev. **36**, 743 (1930).
(76) Gillespie, L. J., and Beattie, J. A.: Phys. Rev. **36**, 1008 (1930).
(77) Gillespie, L. J., and Beattie, J. A.: J. Am. Chem. Soc. **52**, 4239 (1930).

(78) Gillespie, L. J., and Coe, J. R., Jr.: J. Chem. Phys. **1**, 102 (1933).
(79) Gillespie, L. J., and Lurie, E.: J. Am. Chem. Soc. **53**, 2978 (1931).
(80) Gilliland, E. R.: Ind. Eng. Chem. **28**, 212 (1936).
(81) Goff, J. A., Anderson, J. R., and Gratch, S.: Trans. Am. Soc. Heating Ventilating Engrs. **49**, 269 (1943).
(82) Goff, J. A., and Gratch, S.: J. Am. Soc. Heating Ventilating Engrs., Sect. Heating, Piping, Air Conditioning **17**, 334 (1945).
(83) Goranson, R. W.: J. Chem. Phys. **5**, 107 (1937).
(84) Haber, F.: Z. Elektrochem. **21**, 89 (1915).
(85) Hildebrand, J. H.: *Solubility of Non-Electrolytes*. Reinhold Publishing Corporation, New York (1936).
(86) Hirschfelder, J. O., and Roseveare, W. E.: J. Phys. Chem. **43**, 15 (1939).
(87) Kay, W. B.: Ind. Eng. Chem. **28**, 1014 (1936).
(88) Kazarnovskiĭ, Y. S., Simonov, G. B., and Aristov, G. E.: J. Phys. Chem. (U. S. S. R.) **14**, 774 (1940).
(89) Kelso, E. A., and Felsing, W. A.: Ind. Eng. Chem. **34**, 161 (1942).
(90) Keyes, F. G.: Proc. Natl. Acad. Sci. U. S. **3**, 323 (1917).
(91) Keyes, F. G.: Ann. Survey Am. Chem. **2**, 18 (1926–27).
(92) Keyes, F. G.: J. Am. Chem. Soc. **49**, 1393 (1927).
(93) Keyes, F. G.: J. Am. Chem. Soc. **60**, 1761 (1938).
(94) Keyes, F. G.: Nucleus **19**, 205 (1942).
(95) Keyes, F. G., and Burks, H. G., Jr.: J. Am. Chem. Soc. **50**, 1100 (1928).
(96) Kleeman, R. D.: J. Franklin Inst. **206**, 511 (1928).
(97) Krichevsky, I. R.: J. Am. Chem. Soc. **59**, 2933 (1937).
(98) Krichevsky, I. R., and Kasarnowsky, J. J.: Acta Physicochim. U.R.S.S. **10**, 217 (1939).
(99) Kuenan, J. P.: *Theorie der Verdampfung and Verflüssigung von Gemischen*, p. 99. Barth, Leipzig (1906).
(100) Laar, J. J. van: *Sechs Vorträge über das thermodynamische Potential*, p. 82. F. Vieweg und Sohn, Braunschweig (1906).
(101) Laar, J. J. van: Z. physik. Chem. **72**, 723 (1910).
(102) Laar, J. J. van: Z. physik. Chem. **83**, 599 (1913).
(103) Laar, J. J. van: Z. physik. Chem. **A137**, 421 (1928).
(104) Laar, J. J. van: Z. anorg. allgem. Chem. **171**, 42 (1928).
(105) Laar, J. J. van: Z. physik. Chem. **A145**, 207 (1929).
(106) Laar, J. J. van, and Lorenz, R.: Z. anorg. allgem. Chem. **145**, 239 (1925).
(107) Laar, J. J. van, and Lorenz, R.: Z. anorg. allgem. Chem. **146**, 42 (1925).
(108) Larson, A. T.: J. Am. Chem. Soc. **46**, 367 (1924).
(109) Larson, A. T., and Black, C. A.: J. Am. Chem. Soc. **47**, 1015 (1925).
(110) Larson, A. T., and Dodge, R. L.: J. Am. Chem. Soc. **45**, 2918 (1923).
(111) Lashakov, L. M.: Acta. Physicochim. U.R.S.S. **11**, 107 (1939).
(112) Lashakov, L. M.: Acta. Physicochim. U.R.S.S. **11**, 849 (1939).
(113) Lashakov, L. M.: J. Chem. Ind. (U.S.S.R.) **18**, 17 (1941).
(114) Lashakov, L. M.: Khim. Referat. Zhur. **4**, 3 (1941).
(115) Lennard-Jones, J. E.: Proc. Roy. Soc. (London) **A106**, 463 (1924).
(116) Lennard-Jones, J. E.: Physica **4**, 941 (1937).
(117) Lennard-Jones, J. E., and Cook, W. R.: Proc. Roy. Soc. (London) **A115**, 334 (1927).
(118) Lerberghe, G. van: Compt. rend. **181**, 851 (1925).
(119) Lerberghe, G. van: Bull. classe sci., Acad. roy. Belg. [5] **14**, 349 (1928).
(120) Lerberghe, G. van: *Le cacul des fugacités d'une solution*. Marcel Hayez, Bruxelles (1928).
(121) Lerberghe, G. van: Bull. classe sci., Acad. roy. Belg. [5] **16**, 94 (1930).
(122) Lerberghe, G. van, and Schouls, G.: Bull. classe sci., Acad. roy. Belg. [5] **15**, 583 (1929).

(123) LEWIS, G. N., AND RANDALL, M.: *Thermodynamics*, p. 190. McGraw-Hill Book Company, Inc., New York (1923).

(124) Reference 123, p. 221.

(125) Reference 123, pp. 225–7.

(126) LEWIS, W. K.: Oil Gas J. **32**, Oct. 19, p. 14; Oct. 26, p. 20; Nov. 2, p. 18 (1933).

(127) LEWIS, W. K.: Trans. Am. Inst. Mining Met. Engrs. **107**, 11 (1934).

(128) LEWIS, W. K.: Ind. Eng. Chem. **28**, 257 (1936).

(129) LEWIS, W. K., AND KAY, C. W.: Oil Gas J. **32**, Mar. 29, p. 40 (1934).

(130) LEWIS, W. K., AND LUKE, C. D.: Trans. Am. Soc. Mech. Engrs. **54**, PME-8, 55 (1932).

(131) LEWIS, W. K., AND LUKE, C. D.: Ind. Eng. Chem. **25**, 725 (1933).

(132) LORENTZ, H. A.: Ann. Physik **12**, 127, 660 (1881).

(133) LORENZ, R.: *Das Gesetz der chemiken Massenwirkung.* Voss, Leipzig (1927).

(134) LURIE, E., AND GILLESPIE, L. J.: J. Am. Chem. Soc. **49**, 1146 (1927).

(135) MARON, S. H., AND TURNBULL, D.: Ind. Eng. Chem. **33**, 69, 246 (1941).

(136) MARON, S. H., AND TURNBULL, D.: Ind. Eng. Chem. **33**, 408 (1941).

(137) MARON, S. H., AND TURNBULL, D.: Ind. Eng. Chem. **34**, 544 (1942).

(138) MASSON, I., AND DOLLEY, L. G. F.: Proc. Roy. Soc. (London) **A103**, 524 (1923).

(139) MAYER, J. E.: J. Phys. Chem. **43**, 71 (1939).

(140) MERZ, A. R., AND WHITTAKER, C. W.: J. Am. Chem. Soc. **50**, 1522 (1928).

(141) NEWTON, R. H.: Ind. Eng. Chem. **27**, 302 (1935).

(142) NEWTON, R. H., AND DODGE, B. F.: J. Am. Chem. Soc. **56**, 1287 (1934).

(143) NEWTON, R. H., AND DODGE, B. F.: Ind. Eng. Chem. **27**, 577 (1935).

(144) PERKINS, A. J.: J. Chem. Phys. **5**, 180 (1937).

(145) POLLITZER, F., AND STREBEL, E.: Z. physik. Chem. **110**, 768 (1924).

(146) PRIGOGINE, I., AND DEFAY, R.: *Thermodynamique chimique.* Gauthier-Villars, Paris (1944).

(147) RANDALL, M., AND SOSNICK, B.: J. Am. Chem. Soc. **50**, 967 (1928).

(148) ROBINSON, H. M., AND BLISS, H.: Ind. Eng. Chem. **32**, 396 (1940).

(149) RYSSELBERGHE, P. VAN, AND DONDER, T. DE: *The Thermodynamic Theory of Affinity.* Stanford University Press, Stanford University (1936).

(150) SADDINGTON, A. W., AND KRASE, N. W.: J. Am. Chem. Soc. **56**, 353 (1934).

(151) SAGE, B. H., BUDENHOLZER, R. A., AND LACEY, W. N.: Ind. Eng. Chem. **32**, 1262 (1940).

(152) SAGE, B. H., HICKS, B. L., AND LACEY, R. A.: Ind. Eng. Chem. **32**, 1085 (1940).

(153) SAGE, B. H., AND LACEY, W. N.: Ind. Eng. Chem. **30**, 673 (1938).

(154) SAGE, B. H., AND LACEY, W. N.: Ind. Eng. Chem. **31**, 1497 (1939).

(155) SAGE, B. H., AND LACEY, W. N.: Ind. Eng. Chem. **34**, 730 (1942).

(156) SAGE, B. H., LACEY, W. N., AND SCHAAFSMA, J. G.: Ind. Eng. Chem. **27**, 48 (1935).

(157) SAGE, B. H., REAMER, H. H., OLDS, R. H., AND LACEY, W. N.: Ind. Eng. Chem. **34**, 1108 (1942).

(158) SAGE, B. H., SCHAAFSMA, J. G., AND LACEY, W. N.: Ind. Eng. Chem. **26**, 1218 (1934).

(159) SAGE, B. H., WEBSTER, D. C., AND LACEY, W. N.: Ind. Eng. Chem. **29**, 658, 1188 (1937).

(160) SARTORI, G.: Atti X° congr. intern. chim. **2**, 478 (1938).

(161) SCATCHARD, G.: Proc. Natl. Acad. Sci. U. S. **16**, 811 (1930).

(162) SCATCHARD, G.: Chem. Revs. **8**, 321 (1931).

(163) SCATCHARD, G.: J. Am. Chem. Soc. **56**, 995 (1934).

(164) SCATCHARD, G., EPSTEIN, L. F., WARBURTON, J., JR., AND CODY, P. J.: Refrig. Eng. **53**, 413 (1947).

(165) SCATCHARD, G., AND RAYMOND, C. L.: J. Am. Chem. Soc. **60**, 1278 (1938).

(166) SCATCHARD, G., WOOD, S. E., AND MOCHEL, J. M.: J. Phys. Chem. **43**, 119 (1939).

(167) SCATCHARD, G., WOOD, S. E., AND MOCHEL, J. M.: J. Am. Chem. Soc. **61**, 3206 (1939).

(168) SCATCHARD, G., WOOD, S. E., AND MOCHEL, J. M.: J. Am. Chem. Soc. **62**, 712 (1940).

(169) SCATCHARD, G., WOOD, S. E., AND MOCHEL, J. M.: J. Am. Chem. Soc. **68**, 1957 (1946).

(170) Scatchard, G., Wood, S. E., and Mochel, J. M.: J. Am. Chem. Soc. **68**, 1960 (1946).

(171) Selheimer, C. W., Souders, M., Jr., Smith, L. R., and Brown, G. G.: Ind. Eng. Chem. **24**, 515 (1932).

(172) Smith, L. B., Beattie, J. A., and Kay, W. C.: J. Am. Chem. Soc. **59**, 1587 (1937).

(173) Souders, M., Jr., Selheimer, C. W., and Brown, G. G.: Ind. Eng. Chem. **24**, 517 (1932).

(174) Stockmayer, W. H.: J. Chem. Phys. **9**, 398 (1941).

(175) Stockmayer, W. H.: J. Phys. Chem. **9**, 863 (1941).

(176) Su, G.-J., and Chang, C.-H.: J. Am. Chem. Soc. **68**, 1080 (1946).

(177) Su, G.-J., Huang, P.-H., and Chang, Y.-M.: J. Am. Chem. Soc. **68**, 1403 (1946).

(178) Trautz, M., and Emert, O.: Z. anorg. allgem. Chem. **150**, 277 (1926).

(179) Trautz, M., and Narath, A.: Ann. Physik [4] **79**, 637 (1926).

(180) Tunnell, R.: J. Phys. Chem. **35**, 2885 (1931).

(181) Waals, J. D. van der: *Die Continuität des gasförmigen und flüssigen Zustandes.* Barth, Leipzig (1899–1900).

(182) Wettberg, E. F. von, Jr., and Dodge, B. F.: Ind. Eng. Chem. **22**, 1040 (1930).

VOLUMETRIC BEHAVIOR IN SEVERAL GASEOUS HYDROCARBON SOLUTIONS[1]

E. W. HOUGH AND B. H. SAGE

California Institute of Technology, Pasadena, California

Received October 25, 1948

The recent marked accumulation of information concerning the volumetric behavior of hydrocarbons and their mixtures has made it interesting to compare the various methods of prediction of such behavior with the experimental data.

A comparison has been made of volumetric data for the binary mixtures of methane with ethane, propane, and *n*-butane with predictions based on an equation of state for the mixtures. The comparison has been carried out at pressures up to 204 atm. and at temperatures from 38° to 121°C. The results indicate that for regions of low density some of the available methods are adequate for most technical uses. However, at the higher densities corresponding to the higher pressures there still remains the need for additional refinement to permit the use of equations of state in many industrial requirements.

Gaseous solutions are typified in nature by the atmosphere and by natural gas at relatively low pressures. The assumptions of the additivity of volume (18) and of the perfect gas laws can be used to approximate their behavior roughly. The additive volume concept, which is equivalent in part to the assumption of ideal solutions (17), appears to be applicable at reduced temperatures below about two only for low reduced pressures, but at high reduced temperatures for a fair range of reduced pressures. Efforts to correlate the volumetric and phase behavior of mixtures of paraffin hydrocarbons by graphical, tabular, or nomographic means have been notable for their lack of success except for the volumetric generalization proposed by Kay (16) that has come to be known as the "pseudo-critical correlation." This generalization has proven to be most useful for many purposes but is not sufficiently precise to satisfy all requirements nor does it permit direct estimation of phase behavior. These limitations in the relatively simple methods of estimation of the volumetric behavior of gaseous mixtures from a knowledge of the behavior of the components have made necessary the experimental investigation of the behavior of such phases. The work on the volumetric behavior of hydrogen and nitrogen (3, 4, 26) and of air (2, 15) has been supplemented recently by studies of natural gases (20, 25) and of a number of binary mixtures of the paraffin hydrocarbons (10, 20, 21, 22, 24). Some of these data have been employed in the evaluation of the partial volumetric behavior of these hydrocarbons in binary mixtures (23). Such information has been used for the prediction of the compressibility factor of natural gases. Most of the experimental data accumulated serve to indicate that gaseous solutions at elevated pressures deviate materially from the additive volume relationships, especially at low reduced temperatures.

---

[1] Presented at the Symposium on Thermodynamics and the Molecular Structure of Solutions, which was held under the auspices of the Division of Physical and Inorganic Chemistry at the 114th Meeting of the American Chemical Society, Portland, Oregon, September 13 and 14, 1948.

The Beattie–Bridgeman equation (6) has been proposed for application to gaseous mixtures. Values of each coefficient, as applicable to the mixture in the single-phase region, have been taken as derivable from the corresponding coefficients for the pure components. On this basis, which was originally proposed by Beattie (5), the several coefficients of the Beattie–Bridgeman equation may be established by application of the following types of expressions:

$$A_{0_m} = \left(\sum_{i}^{n} n_i A_{0_i}^{1/2} \right.$$

$$B_{0_n} = \left(\sum_{i=1}^{n} n_i B_{0_i} \right)$$

where n_i refers to the mole fraction of the i^{th} component in the mixture.

It has been found that this method of evaluating the coefficients of the Beattie–Bridgeman equation does not yield satisfactory results insofar as the prediction of the volumetric behavior of binary and more complex hydrocarbon mixtures in the gaseous region is concerned. However, some improvements in the evaluation of the coefficients have been suggested (9). Furthermore, the Beattie–Bridgeman equation is not suitable for use at high specific weights. For example, one of the originators has indicated a value of about 4 gram-moles per liter (8) as the limiting density at which the equation is applicable to n-butane. This figure in the case of methane is about 6 gram-moles per liter (7).

In order to overcome the lack of applicability of the Beattie–Bridgeman equation to mixtures in the gaseous region and to materials having high specific weight, Benedict proposed (11, 12) a more complex but nevertheless more accurate expression for the estimation of the volumetric behavior of hydrocarbon mixtures. The expression proposed by Benedict appears to be of sufficient accuracy to permit its use to predict the chemical potential of the components in each of the phases. From such knowledge, the composition of the coexisting phases as a function of temperature and pressure may be estimated. These methods appear feasible of extension to practical engineering use in the estimation of the volumetric and phase behavior of multicomponent hydrocarbon systems.

The extension of such methods to the design of process equipment has been limited by the lack of specific information concerning appropriate values of the coefficients for the components of greater molecular weight than n-butane. Furthermore, the numerical effort associated with the establishment of the composition of the coexisting phases is significant. In the past, this effort prevented the widespread utilization of equations of state for the prediction of the composition of coexisting phases in heterogeneous, multicomponent systems. Recently, with the advent of more effective mechanical or electronic computing devices (see, for example, articles by Aiken (1) and Brainerd (13)), it appears more feasible to solve the requisite simultaneous equations by trial than in the past. Therefore it is believed that the extension of the knowledge of the co-

TABLE 1

Ratio of volumes in the methane–ethane system*

A. Temperature = 37.8°C. (100°F.)

COMPOSITION MOLE FRACTION METHANE	EXPERIMENTAL VOLUME		RATIO OF VOLUMES*
	Liter/gram-mole	Cu. ft./lb. mole	
Pressure = 27.22 atm. (400 lb./sq. in.)			
0.0	0.743	11.90	0.9909
0.1	0.763	12.23	0.9981
0.2	0.783	12.55	1.0015
0.4	0.821	13.15	1.0024
0.6	0.851	13.63	1.0034
0.8	0.875	14.02	1.0045
1.0	0.900	14.41	0.9998
Pressure = 68.05 atm. (1000 lb./sq. in.)			
0.0	0.0964	1.544	0.9954
0.1	0.1410	2.259	0.92†
0.2	0.1828	2.928	1.0058
0.4	0.2485	3.98	1.0064
0.6	0.2890	4.63	1.0045
0.8	0.317	5.08	1.0049
1.0	0.340	5.45	0.9984
Pressure = 204.1 atm. (3000 lb./sq. in.)			
0.0	0.0742	1.189	0.9978
0.1	0.0750	1.202	0.9940
0.2	0.0760	1.217	0.9954
0.4	0.0788	1.262	1.0084
0.6	0.0841	1.347	1.0303
0.8	0.0932	1.493	1.0285
1.0	0.1069	1.713	1.0025

B. Temperature = 121.1°C. (250°F.)

COMPOSITION MOLE FRACTION METHANE	EXPERIMENTAL VOLUME		RATIO OF VOLUMES*
	Liter/gram-mole	Cu. ft./lb. mole	
Pressure = 27.22 atm. (400 lb./sq. in.)			
0.0	1.087	17.41	0.9979
0.2	1.106	17.71	1.0032
0.4	1.124	18.00	1.0037
0.6	1.142	18.29	1.0029
0.8	1.158	18.55	1.0013
1.0	1.173	18.79	0.9985

* Ratio of volume calculated from Benedict equation to experimental volume.

† Value omitted from figure. Curve dotted in the vicinity because exact shape not established.

TABLE 1—*Concluded*

B. Temperature = 121.1°C. (250F.°)—*Concluded*

COMPOSITION MOLE FRACTION METHANE	EXPERIMENTAL VOLUME		RATIO OF VOLUMES*
	Liter/gram-mole	Cu. ft./lb. mole	
Pressure = 68.05 atm. (1000 lb./sq. in.)			
0.0	0.373	5.97	0.9912
0.2	0.395	6.32	1.0132
0.4	0.414	6.63	1.0091
0.6	0.431	6.91	1.0087
0.8	0.447	7.16	1.0065
1.0	0.463	7.41	0.9966
Pressure = 204.1 atm. (3000 lb./sq. in.)			
0.0	0.1089	1.745	0.9951
0.2	0.1174	1.881	1.06†
0.4	0.1261	2.020	1.0173
0.6	0.1352	2.165	1.0235
0.8	0.1446	2.317	1.0172
1.0	0.1546	2.476	1.0008

efficients for the Benedict equation to the paraffinic hydrocarbons of higher molecular weight will be worth while. With the completion of the study of the behavior of binary and ternary systems of olefinic and paraffinic hydrocarbons, it will be possible to develop a sufficient background of information to permit the prediction of the behavior of multicomponent hydrocarbon systems of low and intermediate molecular weight. Such calculations should be capable of accomplishment with the accuracy necessary for nearly any engineering application.

It is the purpose of the present discussion to review the accuracy with which the volumetric behavior of several binary paraffin hydrocarbon systems may be predicted from the available constants for the Benedict equation of state (11, 12). Since no constants for the Benedict equation for paraffin hydrocarbons of higher molecular weight than *n*-butane are available, it has been necessary to limit the systems which may be included in the comparison. In table 1 are presented the values of the ratio of molal volumes estimated from the Benedict equation with those obtained from experimental measurements (22) for the methane–ethane system. The experimentally determined molal volumes have been included in this tabulation. The range of pressures and temperatures includes the greater part of states covered by the experimental investigation. The volume ratios for this system are portrayed graphically in figures 1 and 2. In the figures, dotted lines indicate that the volume ratios have not been established in sufficient detail in the regions indicated to draw the correct curves. If the experimental points in these regions are not in the range of the graph, they are omitted. The agreement between the predicted and experimental values is, for the most part, better for the pure substances than for mixtures, but is reasonably good at nearly all states. It is understood that the data for the methane–ethane system (22) were not employed in the derivation of the constants for methane and ethane in the Benedict equation (11, 12). The standard deviation of the predicted values from the experimental measurements

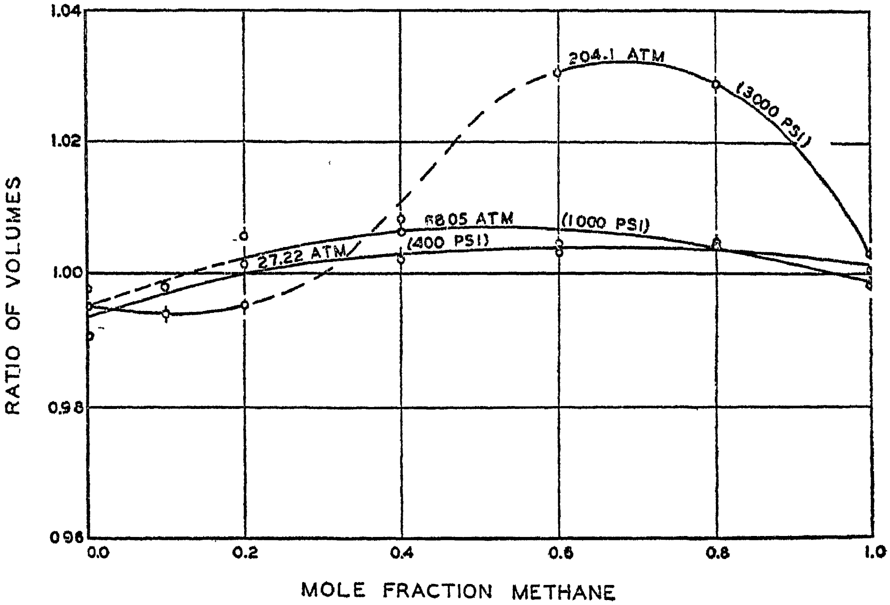


FIG. 1. Ratio of volumes for the methane–ethane system at 37.8°C. (100°F.)

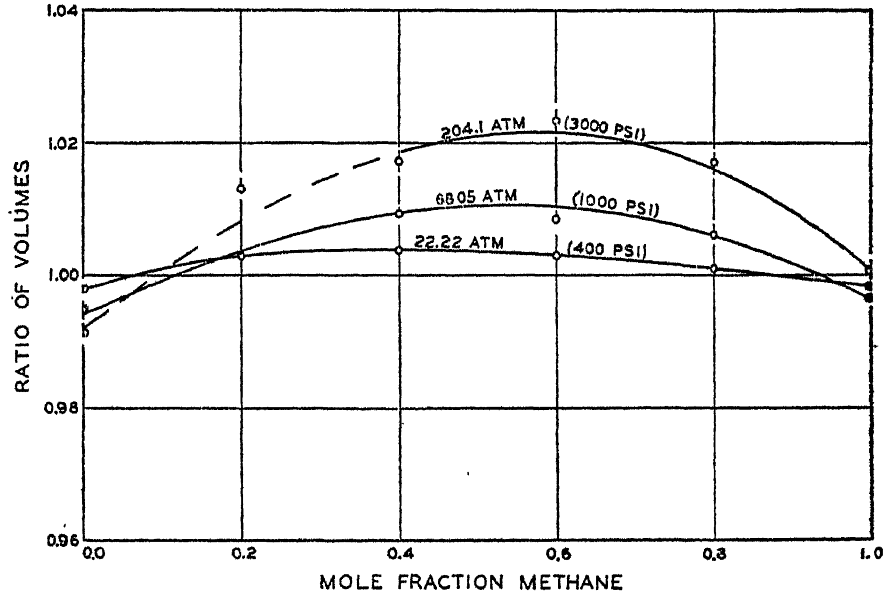


FIG. 2. Ratio of volumes for the methane–ethane system at 121.1°C. (250°F.)

TABLE 2

Ratio of volumes in methane–propane system*

Temperature = 37.8°C. (100°F.)

COMPOSITION MOLE FRACTION METHANE	EXPERIMENTAL VOLUME		RATIO OF VOLUMES*
	Liter/gram-mole	Cu. ft./lb. mole	
Pressure = 27.22 atm. (400 lb./sq. in.)			
0.0	0.0923	1.479	1.021
0.0845†			
0.4419‡			
0.6	0.783	12.55	1.025
0.7	0.823	13.19	1.0086
0.8	0.854	13.68	1.0057
1.0	0.900	14.41	0.9998
Pressure = 68.05 atm. (1000 lb./sq. in.)			
0.0	0.0900	1.441	1.0216
0.2	0.0889	1.424	1.0176
0.3	0.0926	1.484	0.9998
0.3271†			
0.6635‡			
0.7	0.2653	4.25	0.9934
0.8	0.2947	4.72	1.0070
1.0	0.340	5.45	0.9984
Pressure = 204.1 atm. (3000 lb./sq. in.)			
0.0	0.0848	1.358	1.0280
0.2	0.0803	1.287	1.0310
0.3	0.0790	1.266	1.0283
0.6	0.0813	1.302	0.9978
0.7	0.0854	1.368	0.9905
0.8	0.0914	1.464	0.9898
1.0	0.1069	1.713	1.0025

* Ratio of volume calculated from Benedict equation to experimental volume.
† Bubble-point composition.
‡ Dew-point composition.

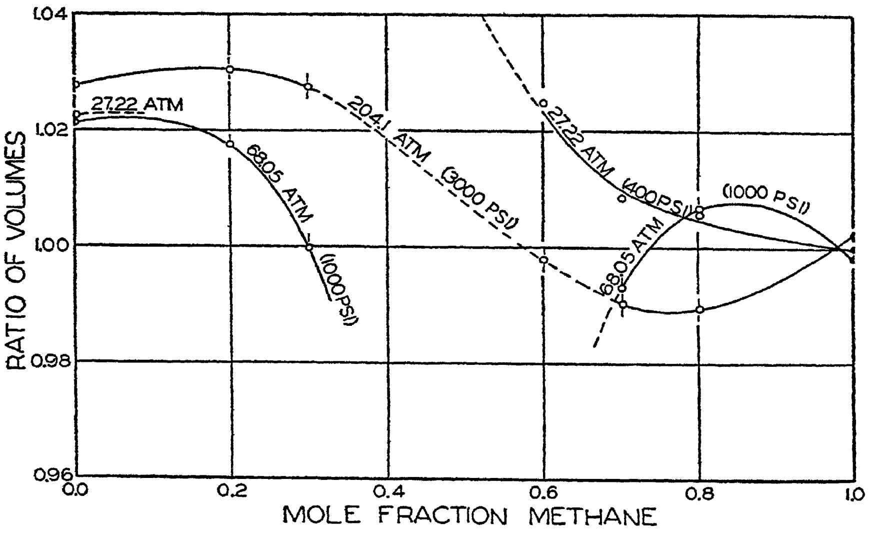


FIG. 3. Ratio of volumes for the methane–propane system at 37.8°C. (100°F.)

TABLE 3

Ratio of volumes in the methane–n-butane system*

A. Temperature = 37.8°C. (100°F.)

COMPOSITION MOLE FRACTION METHANE	EXPERIMENTAL VOLUME		RATIO OF VOLUMES*
	Liter/gram-mole	Cu. ft./lb. mole	
Pressure = 27.22 atm. (400 lb./sq. in.)			
0.0	0.1033	1.655	1.0129
0.1208†			
0.8221‡			
0.9	0.860	13.77	1.0136
0.1	0.900	14.41	0.9998
Pressure = 68.05 atm. (1000 lb./sq. in.)			
0.0	0.1019	1.632	1.0157
0.2	0.0943	1.511	1.0216
0.3	0.0918	1.471	1.0192
0.3172†			
0.8818‡			
0.9	0.312	4.99	0.9974
1.0	0.340	5.45	0.9984
Pressure = 204.1 atm. (3000 lb./sq. in.)			
0.0	0.0984	1.577	1.0204
0.2	0.0895	1.434	1.032
0.3	0.0860	1.378	1.0288
0.6	0.0809	1.296	1.0054
0.8	0.0851	1.364	1.0085
0.9	0.0938	1.503	1.0131
1.0	0.1069	1.713	1.0025

B. Temperature = 121.1°C. (250°F.)

COMPOSITION MOLE FRACTION METHANE	EXPERIMENTAL VOLUME		RATIO OF VOLUMES*
	Liter/gram-mole	Cu. ft./lb. mole	
Pressure = 27.22 atm. (400 lb./sq. in.)			
0.0	0.1358	2.175	0.9952
0.0216†			
0.1417‡			
0.2	0.779	12.48	1.09§
0.4	0.938	15.02	1.0534
0.6	1.039	16.65	1.0350
0.8	1.111	17.80	1.0215
1.0	1.173	18.79	0.9985

* Ratio of volume calculated from Benedict equation to experimental volume.

† Bubble-point composition.

‡ Dew-point composition.

§ Value omitted from figure.

TABLE 3—*Concluded*

B. Temperature = 121.1°C. (250°F.)—*Concluded*

COMPOSITION MOLE FRACTION METHANE	EXPERIMENTAL VOLUME		RATIO OF VOLUMES*
	Liter/gram-mole	Cu. ft./lb. mole	
Pressure = 68.05 atm. (1000 lb./sq. in.)			
0.0	0.1262	2.021	1.0033
0.2	0.1405	2.250	0.9780
0.2192†			
0.4326‡			
0.6	0.358	5.74	1.0192
0.8	0.416	6.66	1.0220
1.0	0.463	7.41	0.9966
Pressure = 204.1 atm. (3000 lb./sq. in.)			
0.0	0.1140	1.826	1.0138
0.2	0.1075	1.722	1.0203
0.4	0.1057	1.693	1.0228
0.6	0.1144	1.832	1.0220
0.8	0.1333	2.135	1.0182
1.0	0.1546	2.476	1.0008

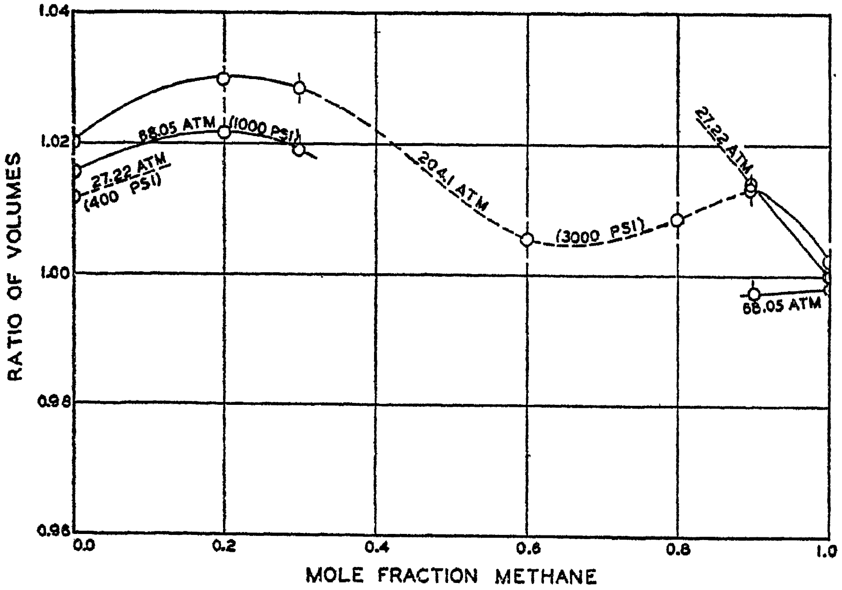


FIG. 4. Ratio of volumes for the methane–n-butane system at 37.8°C. (100°F.)

for the methane–ethane system is about 0.2 per cent for the curve at 27.22 atm. (400 lb./sq. in.) and 121.1°C. (250°F.) and about 2 per cent for the curve at 204.1 atm. (3000 lb./sq. in) and 37.8°C. (100°F.).

A comparison of the ratio of estimated to experimental molal volumes for the

TABLE 4

Ratio of volumes in the methane–ethane system for ideal solutions*

A. Temperature = 37.8°C. (100°F.)

COMPOSITION MOLE FRACTION METHANE	EXPERIMENTAL VOLUME		RATIO OF VOLUMES*
	Liter/gram-mole	Cu. ft./lb. mole	
Pressure = 27.22 atm. (400 lb./sq. in.)			
0.0	0.743	11.90	1.0000
0.2	0.783	12.55	0.9882
0.4	0.821	13.15	0.9810
0.6	0.851	13.63	0.9836
0.8	0.875	14.02	0.9922
1.0	0.900	14.41	1.0000
Pressure = 68.05 atm. (1000 lb./sq. in.)			
0.0	0.0964	1.544	1.0000
0.2	0.1828	2.928	0.7941
0.4	0.2485	3.98	0.7800
0.6	0.2890	4.63	0.8397
0.8	0.317	5.08	0.9198
1.0	0.340	5.45	1.0000
Pressure = 204.1 atm. (3000 lb./sq. in..)			
0.0	0.0742	1.189	1.0000
0.2	0.0760	1.217	1.0635
0.4	0.0788	1.262	1.1084
0.6	0.0841	1.347	1.1161
0.8	0.0932	1.493	1.0773
1.0	0.1069	1.713	1.0000

B. Temperature = 121.1°C. (250°F.)

COMPOSITION MOLE FRACTION METHANE	EXPERIMENTAL VOLUME		RATIO OF VOLUMES*
	Liter/gram-mole	Cu. ft./lb. mole	
Pressure = 27.22 atm. (400 lb./sq. in.)			
0.0	1.087	17.41	1.0000
0.2	1.106	17.71	0.9990
0.4	1.124	18.00	0.9977
0.6	1.142	18.29	0.9975
0.8	1.158	18.55	0.9977
1.0	1.173	18.79	1.0000

* Ratio of volume calculated on the assumption of an ideal solution to experimental volume.

TABLE 4—*Concluded*

B. Temperature = 121.1°C. (250°F.)—*Concluded*

COMPOSITION MOLE FRACTION METHANE	EXPERIMENTAL VOLUME		RATIO OF VOLUMES*
	Liter/gram mole	Cu. ft./lb. mole	
Pressure = 68.05 atm. (1000 lb./sq. in.)			
0.0	0.373	5.97	1.0000
0.2	0.395	6.32	0.9908
0.4	0.414	6.63	0.9878
0.6	0.431	6.91	0.9881
0.8	0.447	7.16	0.9946
1.0	0.463	7.41	1.0000
Pressure = 204.1 atm. (3000 lb./sq. in.)			
0.0	0.1089	1.745	1.0000
0.2	0.1174	1.881	1.0055
0.4	0.1261	2.020	1.0082
0.6	0.1352	2.165	1.0089
0.8	0.1446	2.317	1.0054
1.0	0.1546	2.476	1.0000

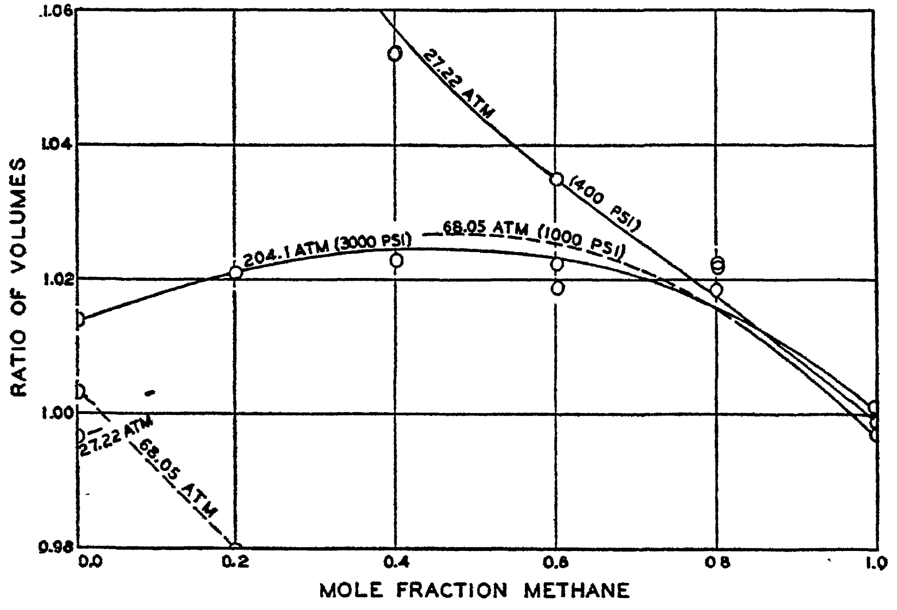


FIG. 5. Ratio of volumes for the methane–*n*-butane system at 121.1°C. (250°F.)

methane–propane system is given in table 2 and figure 3. Similar data for the methane–*n*-butane system are given in table 3 and figures 4 and 5. Fairly satisfactory agreement between the estimated and experimental values was obtained even though in some instances the system was in the liquid phase. In the figures

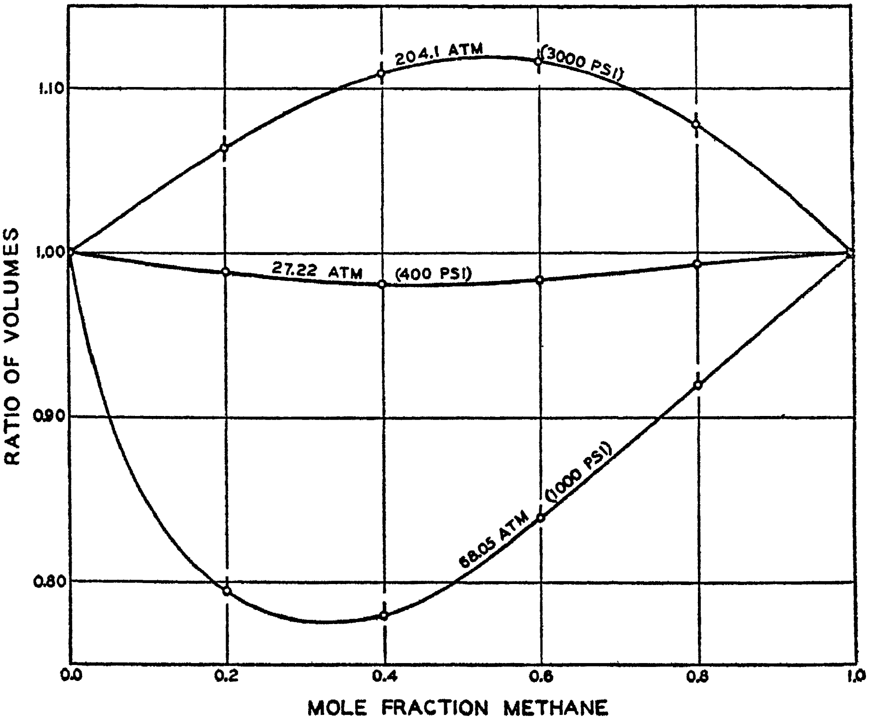


FIG. 6. Ratio of volumes for ideal solution in the methane–ethane system at 37.8°C. (100°F.)

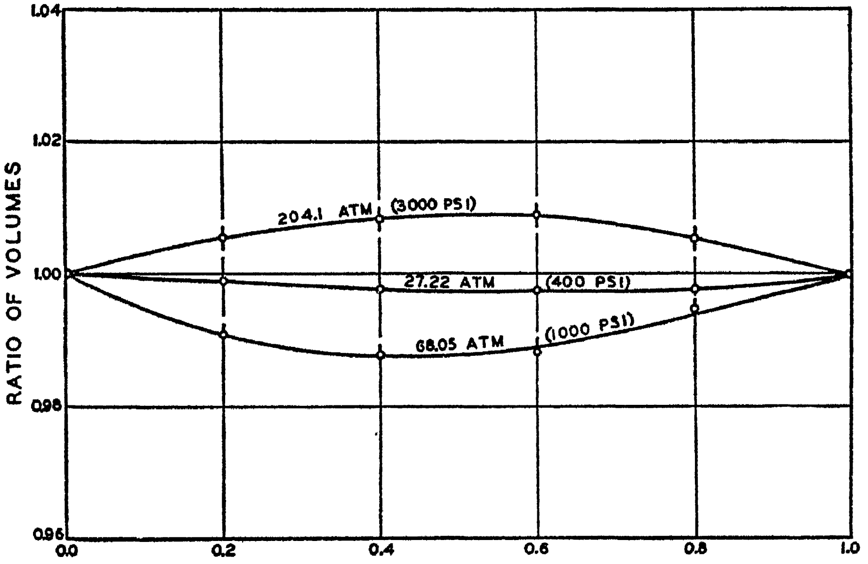


FIG. 7. Ratio of volumes for ideal solution in the methane–ethane system at 121.1°C. (250°F.)

the end of a curve within the boundaries of the graph indicates the presence of a bubble- or a dew-point. As a matter of interest the corresponding experimentally determined molal volumes have been included in the tables along with the bubble-point and dew-point compositions when heterogeneous equilibrium was involved. It is surprising that larger deviations between the predicted and experimental data do not exist at the higher pressures. The assumptions (14) often made in regard to the contribution of the components of a gaseous mixture to the behavior of the system are not applicable at the higher pressures.

A prediction of the volume of one system is made on assumptions of ideal solution because of the use of the method in the process industries. The data for mixtures of methane and ethane are presented in table 4 and in figures 6 and 7. Satisfactory agreement is obtained at lower pressures.

REFERENCES

(1) Aiken, H. H., and Hopper, G. M.: Elec. Eng. **65**, 384, 449, 522 (1946).

(2) Amagat, E. H.: Ann. chim. phys. [6] **29**, 68 (1893).

(3) Bartlett, E. P., Cupples, H. L., and Tremearne, T. H.: J. Am. Chem. Soc. **50**, 1275 (1928).

(4) Bartlett, E. P., Hetherington, H. C., Kvalnes, H. M., and Tremearne, T. H.: J. Am. Chem. Soc. **52**, 1363 (1930).

(5) Beattie, J. A.: J. Am. Chem. Soc. **51**, 19 (1929).

(6) Beattie, J. A., and Bridgeman, O. C.: J. Am. Chem. Soc. **49**, 1665 (1927).

(7) Beattie, J. A., and Bridgeman, O. C.: Proc. Am. Acad. Arts Sci. **63**, 229 (1928).

(8) Beattie, J. A., Simard, G. L., and Gouq-jen, Sui: J. Am. Chem. Soc. **61**, 26 (1939).

(9) Beattie, J. A., and Stockmayer, W. H.: J. Chem. Phys. **10**, 473 (1942).

(10) Beattie, J. A., Stockmayer, W. H., and Ingersoll, H. G.: J. Chem. Phys. **9**, 871 (1941).

(11) Benedict, M., Webb, G. B., and Rubin, L. C.: J. Chem. Phys. **8**, 334 (1940).

(12) Benedict, M., Webb, G. B., and Rubin, L. C.: J. Chem. Phys. **10**, 747 (1942).

(13) Brainerd, J. G., and Sharpless, T. K.: Elec. Eng. **67**, 163 (1948).

(14) Chapman, S., and Cowling, T. G.: *The Mathematical Theory of Non-Uniform Gases*, p. 134. Cambridge University Press, London (1939).

(15) Holborn, L., and Schultze, H.: Ann. Physik **47**, 1089 (1915).

(16) Kay, W. B.: Ind. Eng. Chem. **28**, 1014 (1936).

(17) Lewis, G. N.: J. Am. Chem. Soc. **30**, 668 (1908).

(18) Lewis, G. N., and Randall, M.: *Thermodynamics and the Free Energy of Chemical Substances*, p. 221. McGraw-Hill Book Company, Inc., New York (1923).

(19) Olds, R. H., Sage, B. H., and Lacey, W. N.: Ind. Eng. Chem. **34**, 1008 (1942).

(20) Olds, R. H., Sage, B. H., and Lacey, W. N.: Petroleum Technol. **8**, No. 3 (1945).

(21) Reamer, H. H., Olds, R. H., Sage, B. H., and Lacey, W. N.: Ind. Eng. Chem. **34**, 1526 (1942).

(22) Sage, B. H., and Lacey, W. N.: Ind. Eng. Chem. **31**, 1497 (1939).

(23) Sage, B. H., and Lacey, W. N.: Oil Gas J. **38**, 189 (1939).

(24) Sage, B. H., Reamer, H. H., Olds, R. H., and Lacey, W. N.: Ind. Eng. Chem. **34**, 1108 (1942).

(25) Standing, M. B., and Katz, D. L.: Petroleum Technol. **4**, No. 4 (1941).

(26) Wiebe, R., and Gaddy, V. L.: J. Am. Chem. Soc. **60**, 2300 (1948).

THE TRANSPORT PROPERTIES OF GASES AND GASEOUS MIXTURES. II[1,2]

JOSEPH O. HIRSCHFELDER, R. BYRON BIRD, AND ELLEN L. SPOTZ

Naval Research Laboratory, University of Wisconsin, Madison, Wisconsin

Received October 25, 1948

The intermolecular forces are determined for collisions between both like and unlike molecules from experimental measurements of certain transport properties. From these forces other transport properties and the second virial coefficients for pure gases and gaseous mixtures may be calculated as a function of temperature. In most cases the agreement with experiment is excellent. Only thermal conductivity remains anomalous, and this is due to the transfer of energy from the translational to the internal degrees of freedom. A large number of numerical examples are given for viscosity, self-diffusion, ordinary diffusion, and thermal diffusion for pure substances and mixtures. The extrapolations of our equations to very high temperatures are often more reliable than the direct experimental measurements. The relations are given for multicomponent as well as binary mixtures. Formulae are given for all of the transport properties in terms of the integrals presented in a previous paper (7). This treatment assumes that the energy of interaction of two colliding molecules is the sum of an energy of repulsion varying inversely as the twelfth power of the separation and an energy of attraction varying as the inverse sixth power. The viscosities for seven-component industrial flue and fuel gas mixtures up to 1300°K. are readily calculated and shown to agree with available experimental data.

I. INTRODUCTION

The transport coefficients of simple non-polar gases are dependent upon the forces between the molecules involved. In a previous paper (7) we evaluated all of the collision integrals required in the calculations for the transport coefficients. These integrals were tabulated for eighty-four temperature levels, embracing a very wide range.

The potential energy function for the molecular interaction which was employed is that of the form

$$E(r) = 4\epsilon\,[-(r_0/r)^6 + (r_0/r)^{12}] \tag{1}$$

in which r is the molecular separation, ϵ is the maximum energy of attraction, and r_0 is the collision diameter for low-energy head-on collisions. When the force constants are known, it is possible to predict any of the transport properties by the use of the collision integrals; conversely, the transport coefficients provide a very effective means for the determination of the force constants ϵ and r_0.

---

[1] Presented at the Symposium on Thermodynamics and Molecular Structure of Solutions, which was held under the auspices of the Division of Physical and Inorganic Chemistry at the 114th Meeting of the American Chemical Society, Portland, Oregon, September 13 and 14, 1948.

[2] This work was carried out under Contract NOrd 9938 with the Bureau of Ordnance of the United States Navy.

II. DETERMINATION OF THE INTERMOLECULAR FORCE CONSTANTS

At the present time there is a decided dearth of accurate experimental data both for the transport properties and also for the determination of the second virial coefficients of simple gases and their mixtures. On this account, the values of ϵ and r_0 are by no means uniquely specified and must be altered as more precise experimental data become available. Not until more and better measurements have been obtained will it be reasonable to examine and interpret deviations of the actual energy of interaction from the form shown in equation 1.

TABLE 1A

Force constants between like molecules

FORCE CONSTANTS DETERMINED FROM VISCOSITY DATA			REFERENCE FOR VISCOSITY DATA	FORCE CONSTANTS DETERMINED FROM SECOND VIRIAL COEFFICIENT		REFERENCE FOR SECOND VIRIAL COEFFICIENT
Gas	ϵ/k	r_0		ϵ/k	r_0	
	°K.	Å.		°K.	Å.	
Air	97.0	3.617	(a)	99.2	3.522	(c)
H_2	33.3	2.968	(a)	37.02	2.92	(d, e)
N_2	91.46	3.681	(a)	95.9	3.72	(f)
CO_2	190	3.996	(a)	185	4.57	(g)
N_2O	220	3.879	(a)	189	4.59	(g)
NO	119	3.470	(a)	131	3.17	(g)
CH_4	136.5	3.882	(a)	142.7	3.81	(g)
O_2	113.2	3.433	(a)	117.5	3.58	(f)
CO	110.3	3.590	(a)	95.33	3.65	(e)
A	124.0	3.418	(b)	119.5	3.41	(h)
Ne	35.7	2.80	(b)	35.7	2.74	(h)
He	6.03	2.70	(b)	6.03	2.63	(h)

(a) H. L. Johnston and K. E. McCloskey: J. Phys. Chem. **44,** 1038 (1939).

(b) H. L. Johnston and E. R. Grilly: J. Phys. Chem. **46,** 948 (1942).

(c) L. Holborn and H. Schultze: Ann. Physik **47,** 1089 (1915).

(d) J. DeBoer and A. Michels: Physica **5,** 945 (1938); **6,** 97 (1939).

(e) J. Corner: Proc. Roy. Soc. (London) **58,** 737 (1946).

(f) J. O. Hirschfelder and W. E. Roseveare: J. Phys. Chem. **43,** 15 (1939).

(g) J. O. Hirschfelder, F. T. McClure, C. F. Curtiss, and D. W. Osborne: NDRC Report A116.

(h) R. A. Buckingham: Proc. Roy. Soc. (London) **A168,** 264 (1938).

We have already discussed (7) the calculation of force constants between pairs of similar molecules from viscosity data for pure gases. In table 1 are tabulated values of ϵ/k and r_0 so obtained. There are also shown some force constants which are estimated from critical and boiling-point data. These values were obtained using the following approximate relations (which follow from the Lennard-Jones and Devonshire theory of gases and liquids (11; also 2)):

$$\epsilon/k \doteq 0.75T_c \tag{2}$$

$$\epsilon/k \doteq 1.39T_b \tag{3}$$

$$r_0 \doteq (8.33 \times 10^7)(V_c)^{1/3} \tag{4}$$

TABLE 1B

Force constants between like molecules

GAS	r_0 FROM VISCOSITY DATA	ϵ/k FROM VISCOSITY DATA	ϵ/k FROM BOILING POINT	ϵ/k FROM CRITICAL TEMPERATURE	r_0 FROM CRITICAL VOLUME	REFERENCE FOR VISCOSITY DATA
	Å.	°K.	°K.	°K.	Å.	
Xe	4.051	230	228	192	4.04	(i)
C_2N_2	4.38	339	352	301		(i)
AsH_3	4.06	281	303			(i)
CH_2Cl_2	4.759	406	435			(j)
Kr	3.61	190	169	158	3.96	(i)
Cl_2	4.115	357	332	313	4.15	(j)
Br_2	3.859	400	461	431		(k)
COS	4.13	335	310	284		(i)
I_2	4.982	550	637	620		(j)
CCl_4	5.881	327	487	417	5.42	(i)
HgI_2	5.625	698	872			(j)
$SnBr_4$	6.666	465	660			(j)
CS_2	4.438	488	445	410		(p)
$HgBr_2$	5.414	530	827			(j)
CH_3Cl	3.375	855	346	312	4.29	(j)
$CHCl_3$	5.430	327	581	402	5.11	(i)
HCl	3.305	360	261	243	3.69	(m)
SO_2	4.290	252	366	323	4.15	(n)
$SnCl_4$	4.540	1550	538	444	5.88	(j)
$CH\equiv CH$	4.221	185	263	232	4.02	(l)
$CH_2=CH_2$	4.232	205	235	212	4.19	(i)
C_2H_6	4.418	230	257	229	4.36	(i)
C_3H_8	5.061	254	321	278		(i)
$n\text{-}C_4H_{10}$	4.997	410	379	320		(l)
$i\text{-}C_4H_{10}$	5.341	313	366	305		(l)
$n\text{-}C_5H_{12}$	5.769	345	431	353	5.64	(p)
$n\text{-}C_6H_{14}$	5.909	413	475	381	5.97	(p)
$n\text{-}C_8H_{18}$	7.451	320	555	427	6.56	(i)
$n\text{-}C_9H_{20}$	8.448	240	589			(i)
Cyclohexane	6.093	324	491	416	5.65	(p)
C_6H_6	5.270	440	491	442	5.30	(i)
CH_3OH	3.585	507	470	385	4.08	(p)
C_2H_5OH	4.455	391	489	387	4.59	(p)

(i) Landolt-Börnstein: *Physikalisch-Chemische Tabellen.*

(j) H. Braune and R. Linke: Z. physik. Chem. **A148**, 195 (1930).

(k) A. O. Rankine: Proc. Roy. Soc. (London) **88**, 582 (1913).

(l) T. Titani: Bull. Chem. Soc. Japan **5**, 98 (1930).

(m) M. Trautz and A. Narath: Ann. Physik **79**, 637 (1926).

(n) M. Trautz and R. Zink: Ann. Physik **7**, 427 (1930).

(p) T. Titani: Bull. Chem. Soc. Japan **8**, 255 (1933).

in which k is Boltzmann's constant, r_0 is the low-velocity collision diameter expressed in Ångströms, T_c is the critical temperature, V_c is the critical volume in cubic centimeters, and T_b is the boiling point. By means of these approximate relations it is possible to estimate the transport properties as well as the second virial coefficients without any direct experimental data.

The values given in table 1, which are calculated from the viscosity data, should be quite good for those spherical non-polar gases whose interaction can be represented fairly well by the above-mentioned potential function. There are included in the table, however, a number of gases which certainly do not meet these qualifications. While we realize that the calculation of these values is not strictly justified on the bases of the polarity and/or the non-spherical shape of many of the molecules listed, we have set forth these force constants in the hope that they may be of some use in the estimation of the various transport properties until the theory can be extended to include these special cases. Also, these figures give some qualitative ideas as to the size of the molecules and the strength of the interaction forces.

The force constants determined from viscosity seem unreasonable for carbon tetrachloride, stannic bromide, stannic chloride, methyl chloride, carbon disulfide, mercuric iodide, and mercuric bromide. In the case of the tetrahalides, this is somewhat surprising because of their symmetric nature. The carbon disulfide discrepancy can be due to its cigar shape. The mercury halides are quite polar.

For example, the ϵ/k value for stannic chloride which was calculated from viscosity data is far greater than that estimated from the formulae in equations 2 and 3. This discrepancy might suggest that stannic chloride molecules are associated in the vapor phase or are polar. Joel Hildebrand suggests that the chlorine atoms might stand out from the central tin atom like the quills of a porcupine. As a result, there is a large interchange of energy between translation and rotation during collisions. Chapman and Cowling show that the viscosity for such "perfectly rough" spheres may be anomalously large by as much as 14 per cent.

It is not surprising that the force constants for the hydrocarbons do not behave properly. In order to fit their second virial coefficients, it is necessary to use a form for the intermolecular energy involving three adjustable constants rather than two (9). We are trying to evaluate the collision integrals for this three-constant model and hope that it will prove suitable for the hydrocarbons.

There are some substances for which our force constants could not be obtained from the experimental viscosity data. They are listed in table 3, together with their Sutherland's constants and the temperature index of their viscosity. For these substances there is no possible choice of ϵ/k and r_0 which would lead to the observed temperature dependence. Their Sutherland's constants divided by the temperature or their temperature exponent, s, are generally larger than any values which can be interpreted in terms of our equations. This may be seen by comparing the values in table 3 with those in table 5. Obviously, our simple model of the intermolecular potential (designed for spherical non-polar molecules) does not apply in these cases. These molecules fall into the following classes:

A. Polar molecules: H_2O, NH_3, HBr, HCN, HI, $HgCl_2$

The energy of interaction of polar molecules is quite different from that of non-polar molecules. Stockmayer (14) has considered the interaction of two polar spherical molecules. He finds that, in addition to the interaction energy which we have assumed (equation 1), there is a term proportional to the product of the dipole moments, inversely proportional to the cube of the separation, and dependent on the orientation of the molecules. With this Stockmayer potential it is possible to fit the second virial coefficients for simple polar molecules with high precision. In order to predict the transport properties for polar molecules, the collision integrals should be evaluated for this same potential. The work involved is stupendous, but sooner or later it will have to be done.

TABLE 2

Force constants between unlike molecules

GAS PAIR	EXPERIMENTAL D_{12} MEASUREMENT	TEMPERATURE FOR D_{12}	ARITHMETIC MEAN r_{12} FROM VISCOSITY	GEOMETRIC MEAN ϵ_{12}/k FROM VISCOSITY	ϵ_{12}/k FROM D_{12}	ϵ_{12}/k FROM BINARY VISCOSITIES	REFERENCE FOR D_{12} DATA
	$cm.^2\ sec.^{-1}$	°K.	Å.	°K.	°K.	°K.	
He–A	0.641	273.2	3.059	27.3	24.5	30.1	(a)
H_2–O_2	0.697	273.2	3.201	61.4	58.2	59.8	(a)
H_2–N_2	0.674	273.2	3.325	55.2	47.8	60.4	(a)
	0.743	288.2			46.1		(b)
	0.76	293.2			47.7		(c)
H_2–CH_4	0.625	273.2	3.395	67.4	63.8	53.6	(a)
O_2–N_2	0.181	273.2	3.557	102	90.7	98.5	(a)
	0.22	293.2			67.7		(c)
O_2–CO	0.185	273.2	3.512	112	91.9	116	(a)
CO–N_2	0.192	273.2	3.636	100	68.2	104	(a)
CO_2–N_2O	0.096	273.2	3.938	204	183	190	(a)
H_2–CO	0.651	273.2	3.279	60.6	65.0		(a)
H_2–CO_2	0.550	273.2	3.482	79.5	76.0		(a)
	0.619	288.2			63.8		(b)
	0.60	293.2			88.4		(c)
H_2–N_2O	0.535	273.2	3.424	85.6	96.2		(a)
O_2–CO_2	0.139	273.2	3.715	147	115		(a)
	0.16	293.2			110		(c)
CO–CO_2	0.137	273.2	3.793	145	120		(a)
CO_2–N_2	0.144	273.2	3.839	132	93.3		(a)
	0.158	288.2			93.9		(b)
	0.16	293.2			100		(c)
CO_2–CH_4	0.153	273.2	3.909	161	126		(a)
H_2–A	0.77	293.2	3.193	64.3	64.1		(c)
A–N_2	0.20	293.2	3.550	107	83.8		(c)
A–O_2	0.20	293.2	3.426	119	95.2		(c)
A–CO_2	0.14	293.2	3.707	154	142		(c)
He–Ne			2.75	14.7		12.7	
Ne–A			3.11	66.5		68.9	

(a) S. Chapman and T. G. Cowling: *Mathematical Theory of Non-Uniform Gases*, Cambridge University Press, Teddington (1939).

(b) L. E. Boardman and N. E. Wild: Proc Roy. Soc. (London) **A162, 511** (1937).

(c) L. Waldman: Naturwissenschaften **32**, 223 (1944).

The energy of interaction between polar and non-polar molecules is of the same form as the energy of interaction between two non-polar molecules. It is a surprising fact that the effect of the permanent dipole moment of the polar molecule is negligible in attracting the non-polar molecule (12). However, on this account we can treat such collisions as though they were between two non-polar molecules. Thus, in a forthcoming paper we shall consider the properties of gaseous mixtures in which one component is polar.

TABLE 3

Viscosity of anomalous gases

$$\eta \times 10^7 = c' T^s$$
$$\eta \times 10^7 = k_S T^{1/2}/[1 + (S/T)]$$

GAS	TEMPERATURE RANGE	c'	s	k_S	S	(S/T) SUTHERLAND CONSTANT DIVIDED BY TEMPERATURE IN MIDDLE OF RANGE	REFERENCE FOR VISCOSITY DATA
	°K.						
H_2O	300–400	2.039	1.079	140.2	459.4	1.313	(a)
	500–600	1.227	1.164	235.8	1051	1.911	
	600–700	1.598	1.123	244.4	1108	1.705	
NH_3	300–400	1.203	1.181	202.7	740.7	2.116	(a)
	500–600	2.576	1.053	189.4	684.3	1.244	
	600–700	5.207	0.9427	164.7	518.7	0.7980	
HI	300–400	6.889	0.9837	221.9	312.9	0.8940	(b)
	400–500	10.42	0.9152	229.4	334.4	0.7431	
HBr	300–400	5.004	1.040	245.3	376.2	1.075	(a)
HCN	300–400	0.7443	1.215	166.3	836.2	2.389	(a)
	500–600	1.131	1.144	185.4	999.1	1.817	
$HgCl_2$	500–600	1.521	1.180	351.6	1191	2.165	(a)
	600–700	4.841	1.000	248.1	656.5	1.010	
	750–850	3.406	1.057	314.1	982.8	1.229	
Zn	850–920	1.384	1.237	757.1	2374	2.682	(b)
	920–950	18.68	0.8554	338.9	557.0	0.5957	
Cd	750–850	5.390	1.048	475.5	1011	1.264	(b)
	850–900	9.836	0.9592	406.0	734.2	0.8391	
Hg	500–600	7.488	1.039	486.1	640.1	1.164	(b)
	700–800	8.793	1.018	552.0	776.4	1.035	
	800–900	8.827	1.018	595.7	892.6	1.050	
n-C_7H_{16}	350–450	2.163	0.9789	72.73	363.2	0.9080	(b)
	450–550	0.6715	1.172	133.1	1022	2.044	
F_2	50–150	3.237	1.170	244.7	246.9	2.469	(c)
	200–300	13.52	0.8793	170.7	139.3	0.5572	

(a) H. Braune and R. Linke: Z. physik. Chem. **A148**, 195 (1930).
(b) Landolt-Börnstein: *Physikalisch-Chemische Tabellen.*
(c) E. Kanda: Bull. Chem. Soc. Japan **12**, 463 (1937).

B. Metal vapors and valence-unsaturated molecules: Hg, Cd, and Zn

Valence-unsaturated molecules behave in a rather anomalous fashion. Sometimes their collisions follow one potential energy curve and sometimes another. At large separations, their van der Waals energy of attraction is perfectly normal. However, at shorter distances the energy of interaction depends very largely on the particular electronic state, which corresponds to the collision complex. According to quantum mechanics, there are well defined *a priori* probabilities that a collision will result in a particular electronic state.

The collision between two hydrogen atoms should serve to clarify this complicated situation. At large separations, the van der Waals energy of attraction is perfectly normal. However, at shorter distances the energy of interaction depends upon the spin orientations of the atoms (and the electronic state is determined by the spin orientations). In an arbitrary collision between two hydrogen atoms there is a chance of one in four that the collision will correspond to a singlet electronic state and a chance of three in four that the system will be in a triplet electronic state. The singlet corresponds to the normal H_2 molecule and has the force constants: $\epsilon/k = 51{,}000$°K. and $r_0 = 0.5$ Å. The triplet state corresponds to the lowest excited state of the hydrogen molecule—a state in which the molecules repel each other except at the largest separations. The triplet state has the force constants:[3] $\epsilon/k = 3.8$°K. and $r_0 = 3.5$ Å. Thus, in collisions between hydrogen atoms it is expected that one in four collisions will follow the potential curve corresponding to $\epsilon/k = 51{,}000$°K. and $r_0 = 0.5$ Å., while the other three out of four collisions will follow the potential curve corresponding to $\epsilon/k = 3.8$°K. and $r_0 = 3.5$ Å.

The metallic vapors mercury, cadmium, and zinc are not valence saturated, and it is to be expected that their collisions should be anomalous in the same sense as those of the hydrogen atoms. These metal vapor atoms are initially in the $^1S_0$ state, but their valence electrons are easily promoted during the course of collisions.

Hildebrand, Wakeham, and Boyd (6) have determined the interaction potential between mercury atoms from a study of the experimental Menke distribution function for liquid mercury. They found that the mercury atoms are so "squishy" that it is necessary to use the inverse ninth rather than the inverse twelfth power to describe their energy of repulsion. This use of the smaller power might correspond to our view that collisions between mercury atoms are a statistical mixture of two sorts—those with large ϵ/k and small r_0 (corresponding to the ground state of Hg_2) and those with small ϵ/k and large r_0 (corresponding to an excited state of Hg_2). If we try to fit Hildebrand, Wakeham, and Boyd's potential with an inverse twelfth power, we obtain the force constants $\epsilon/k = 1522$ and $r_0 = 2.5$ Å. If we insert these force constants into equation 6 we calculate values for the viscosity of mercury as shown in table 4. There is a considerable discrepancy between the calculated and the observed values of the viscosity. It will be interesting to find out whether this discrepancy is due to the two types of collisions or just to the natural "squishiness" of these particular atoms.

---

[3] Rough quantum-mechanical calculations by J. O. Hirschfelder (unpublished).

C. Cigar-shaped molecules: n-heptane

The transport properties are not very sensitive to the shape of the molecules. However, if the ratio of the length to the diameter becomes too great, all of the physical properties are affected. Thus, the temperature dependence of the viscosity of *n*-heptane is greater than would be expected for spherical molecules.

D. Fluorine

We cannot guess why the temperature dependence of the viscosity of fluorine should be anomalous. Perhaps the experimental data should be checked. If it is substantiated, we would have to conclude that the forces between two fluorine molecules are of a different nature from those for any of the other halogens. This should have a bearing on fluorine chemistry.

TABLE 4

Viscosity of mercury

T	$\eta \times 10^7$ (CALCULATED)	$\eta \times 10^7$ (EXPERIMENTAL)*
°K.		
491	4945	4709
603	5943	5831
769	7466	7610
883	8551	8802

* Landolt-Börnstein: *Physikalisch-Chemische Tabellen.*

The force constants between pairs of dissimilar molecules could very easily be calculated from the coefficient of diffusion as a function of temperature, if these data were available. The expression for the coefficient of diffusion is independent of the force constants for the individual components; hence this particular transport property is exceedingly well-adapted to the calculation of the ϵ_{12} and r_{12} for pairs of unlike molecules. The *International Critical Tables* give values of the diffusion constant for a number of pairs of gases at a single temperature. Using these figures and making the fairly reasonable assumption that the collision diameter, r_{12}, is given by:

$$r_{12} = (r_1 + r_2)/2 \qquad (5)$$

(here r_1 and r_2 are the individual collision diameters), we have calculated the values of ϵ_{12}/k for these molecule pairs. These results are displayed in table 2. It will be noticed that in many cases the values of ϵ_{12}/k can be fairly well approximated by the geometric mean of the ϵ_1/k and ϵ_2/k for the components involved. The values of ϵ_{12}/k determined from diffusion coefficients lead to values of the viscosity of mixtures which agree well with experimental data.

In table 2 we have omitted the consideration of D_{12} for D_2 diffusing through

H_2. The recent value of D_{12} given by Waldman (20) agrees perfectly with our expectation on the basis of the force constants already obtained from the viscosity of H_2. An older experimental value of D_{12} quoted by Chapman and Cowling was considerably in error.

III. THE COEFFICIENT OF VISCOSITY

In this and the following sections, the formulae for the various transport properties will be set forth and their application will be discussed. The formulae for the transport coefficients are essentially those of Chapman and Cowling (1) and are written here for convenience in terms of the transport integrals which are tabulated in table 1 of our preceding publication (7).

The coefficient of viscosity for a single gas is given by the relation:

$$\eta \times 10^7 = 266.93(MT)^{1/2} r_0^{-2} \mathrm{V}/[W^{(2)}(2)] \qquad (6)$$

where M is the molecular weight, T is the temperature in degrees K., r_0 is the low-energy collision diameter expressed in Ångströms, and V and $W^{(2)}(2)$ are slowly varying functions of kT/ϵ previously tabulated. We have already shown how this equation may be used to predict values of the viscosity. And to obtain the intermolecular force constants equation 6 is simply written for values of viscosities at two temperatures, the resulting two equations then being solved simultaneously for r_0 and ϵ.

The temperature dependence of viscosity is often expressed in terms of a Sutherland constant, S:

$$\eta = k_S T^{1/2}/[1 + (S/T)] \qquad (7)$$

Both k_S and S are supposed to be constants characteristic of the gas. Unfortunately, our relations show that this cannot be true over any large temperature range. If we require that at a particular temperature η and $d\eta/dT$ as determined from equation 7 agree with that determined from equation 6, it follows that S/T is the following function of ϵ/kT:

$$S/T = -(4 - \phi_1 - \phi_2 + \phi_3)/(3 - \phi_1 - \phi_2 + \phi_3) \qquad (8)$$

in which:

$$\phi_1 = W^{(2)}(3)/W^{(2)}(2) \qquad (9)$$

$$\phi_2 = 2\left[\frac{\mathrm{V} - 1}{\mathrm{V}}\right]\left[\frac{28W^{(2)}(2) - 17W^{(2)}(3) + 2W^{(2)}(4)}{7W^{(2)}(2) - 2W^{(2)}(3)}\right] \qquad (10$$

$$= 4\left[\frac{(\mathrm{V} - 1)^2}{\mathrm{V}}\right]\left[\frac{\begin{gathered}[\tfrac{77}{8}W^{(2)}(2) + W^{(2)}(4)][4W^{(2)}(2) - W^{(2)}(3)] \\ + W^{(2)}(2)[6W^{(2)}(4) - W^{(2)}(5)] \\ - 2W^{(2)}(3)[5W^{(2)}(3) - W^{(2)}(4)]\end{gathered}}{7W^{(2)}(2) - 2W^{(2)}(3)}\right] \qquad (11)$$

A plot of S/T as a function of kT/ϵ is shown in figure 1 and values are given in table 5. This result is useful, since ϵ/k can be determined immediately from tabulated values of S. The double values of ϵ/kT for a given value of S need cause no concern, since we usually know whether a gas is above or below the critical point ($\epsilon/kT_c = 0.75$).

For some purposes it is more convenient to express the viscosity in terms of an equation of the form:

$$\eta = k'T^s \tag{12}$$

TABLE 5

Temperature dependence of Sutherland's constant and the temperature exponent, s

$$\eta = k_S T^{1/2}/[1 + (S/T)]$$
$$\eta = k'T^s$$

kT/ϵ	S/T	s	kT/ϵ	S/T	s
0.30	0.5538	0.8564	3	0.3669	0.7684
0.50	0.8786	0.9677	4	0.2829	0.7205
0.75	1.083	1.0200	5	0.2453	0.6970
1.00	1.012	1.0030	10	0.1857	0.6566
1.25	0.8660	0.9641	50	0.1697	0.6451
1.50	0.7443	0.9267	100	0.1681	0.6439
2.00	0.5466	0.8534	400	0.1680	0.6438
2.50	0.4331	0.8022			

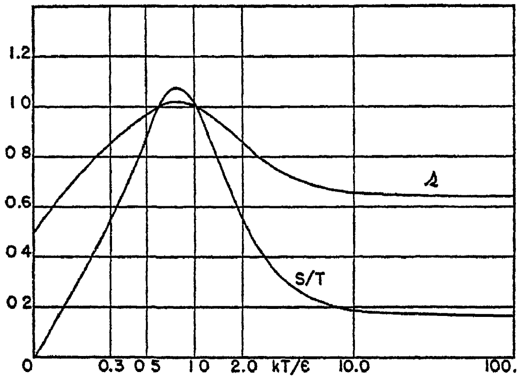


FIG. 1. Sutherland's constant (equation 7) divided by the temperature, S/T, and the temperature exponent, s (equation 12), plotted against kT/ϵ.

The exponent, s, may be determined as a function of ϵ/kT in a manner quite similar to the method used for the Sutherland constant above. It is not difficult to show that

$$s = (9/2) - \phi_1 - \phi_2 + \phi_3 \tag{13}$$

Values of s so determined are shown in table 5, and graphically in figure 1.

For the binary mixture, the formula for the first approximation to the viscosity has already become rather complicated:

$$\eta_{12} \times 10^7 = \frac{R_1 + R_2 + R_3 + (E/H_1) + (E/H_2)}{(R_1/H_1) + (R_2/H_2) + (E/H_1H_2) + (R_4/E)} \tag{14}$$

in which

$$R_1 = (x_1/x_2)(2/3 + AM_1/M_2) \tag{15}$$

$$R_2 = (x_2/x_1)(2/3 + AM_2/M_1) \tag{16}$$

$$R_3 = 2[(2/3) - A] \tag{17}$$

$$R_4 = 2A(M_1 + M_2)^2/3M_1M_2 \tag{18}$$

$$H_i = 266.93(M_iT)^{1/2}(r_i)^{-2}/(W^{(2)}(2; kT/\epsilon_i)) \qquad (i = 1, 2) \tag{19}$$

$$E = 37.75[(M_1 + M_2)^3T/(M_1M_2)]^{1/2}(r_{12})^{-2}/(W^{(1)}(1; kT/\epsilon_{12})) \tag{20}$$

Here x_i is the mole fraction of the i^{th} component; M_i is the molecular weight of the i^{th} component; the r_i and r_{12} are the low-velocity collision diameters measured in Ångströms; A is a function of the collision integrals and is tabulated as a function of kT/ϵ in our previous paper. The H_i are simply the first approximations to the viscosity of the i^{th} component.

Using the force constants between like molecules which were determined from viscosity data and those between dissimilar molecules which were determined from diffusion, we have calculated the viscosity of several binary mixtures. The diffusion data used to get the ϵ_{12}/k are those values tabulated in Chapman and Cowling (1). The experimental binary viscosity data listed are from Landolt–Börnstein, with the single exception of the H_2–N_2 data (18). The results of these computations are displayed in table 6 along with the experimental figures. It will be observed that the agreement is excellent in most cases. This is a nice illustration of the use of force constants obtained from one transport property in the calculation of another of the transport coefficients.

The determination of force constants between dissimilar molecules from viscosity data of binary mixtures is not too difficult in principle. However, the existing data are apparently not sufficiently good to give worthwhile results.

The viscosity of multicomponent gas mixtures has been worked out by C. F. Curtiss and J. O. Hirschfelder (3). For a mixture of f components, the generalized viscosity may be expressed as follows:

$$\eta_{123\ldots f} \times 10^7 = \frac{\begin{array}{ccccc} J_{11} & J_{12} & J_{13}\cdots J_{1f} & 1 \\ J_{12} & J_{22} & J_{23}\cdots J_{2f} & 1 \\ J_{13} & J_{23} & J_{33}\cdots J_{3f} & 1 \\ \\ \\ J_{1f} & J_{2f} & J_{3f}\cdots J_{ff} & 1 \\ 1 & 1 & 1\cdots 1 & 0 \end{array}}{|J_{ij}|} \tag{21}$$

in which $|J_{ij}|$ is the f^{th}-order determinant of the J_{ij}, the J_{ij} (for $i \neq j$) and J_{ii} being given by:

$$J_{ij} = +A_{ij} - B_{ij} \tag{22}$$

TABLE 6

Viscosity of binary mixtures

VISCOSITY H_2–He

T	%H_2	0 0	55.20	60.69	69.18	100.0
°K.						
293	EX* . .	1974	1317	1252	1166	875
	CG* ...	1926	1318	1261	1175	876
373	EX . .	2320	1551	1478	1383	1029
	CG.. ..	2244	1536	1470	1370	1021
473	EX... .	2715	1817	1728	1619	1211
	CG.....	2615	1790	1714	1596	1185
523	EX	2903	1939	1852	1732	1296
	CG.....	2790	1910	1828	1702	1265

VISCOSITY H_2–C_2H_6

T	%H_2..... .	0.0	45.00	85.15	100.0
°K.					
293	EX	909	987	993	876
	CG	909	986	1005	876
373	EX	1142	1208	1189	1033
	CG	1139	1213	1197	1021
473	EX	1409	1467	1412	1213
	CG	1401	1469	1411	1185
523	EX	1526	1583	1511	1296
	CG	1522	1588	1513	1265

VISCOSITY CH_4–C_2H_6

T	%CH_4	0.0	19.03	48.74	81.16	100.0
°K.						
293	EX....	909	938	986	1046	1087
	CG.....	909	939	991	1053	1092
373	EX...	1142	1174	1226	1288	1331
	CG.....	1139	1171	1226	1289	1327
473	EX....	1409	1442	1496	1562	1603
	CG.....	1401	1435	1490	1552	1588
523	EX.....	1526	1560	1614	1682	1725
	CG.....	1522	1556	1611	1674	1709

VISCOSITY C_2H_6–C_3H_8

T	%C_2H_6...	0.0	15.26	25.63	43.27	100.0
°K.						
293	EX	801	815	828	842	909
	CG	801	814	824	824	909
373	EX	1008	1025	1039	1058	1142
	CG	1008	1025	1037	1058	1139
473	EX	1253	1272	1298	1313	1409
	CG	1248	1268	1282	1307	1401
523	EX	1363	1382	1401	1425	1526
	CG	1360	1381	1396	1422	1522

VISCOSITY CH_4–C_3H_8

T	%CH_4......	0.0	16.59	36.17	63.16	100.0
°K.						
293	EX.....	801	831	878	948	1087
	CG.....	801	835	881	957	1092
373	EX.....	1008	1042	1101	1182	1331
	CG	1008	1047	1099	1183	1327
473	EX . ..	1253	1291	1355	1441	1603
	CG.....	1248	1290	1347	1438	1588
523	EX	1363	1403	1465	1553	1725
	CG.....	1360	1404	1463	1556	1709

VISCOSITY NO–N_2

T	%NO .	0.0	26.74	58.37	69.48	100.0
°K.						
293	EX	1747	1778	1827	1833	1882
	CG	1747	1787	1833	1849	1894
373	EX	2084	2132	2209	2222	2272
	CG	2086	2138	2201	2223	2284

---

* EX = experimental data; CG = calculated using a geometric mean ϵ/k; CD = calculated using an ϵ/k calculated from diffusion.

TABLE 6—*Continued*

VISCOSITY H_2–CO_2							VISCOSITY H_2–O_2						
T	%H_2. .	0.0	19.93	41.29	78.50	100.0	%H_2	0.0	18.35	39 45	60.30	78.08	100.0
°*K*.													
300	EX* ...	1493	1501	1506	1370	891		2057	2019	1925	1784	1494	889
	CG* ..	1493	1507	1508	1372	889		2064	2021	1934	1774	1524	889
	CD*....	1493	1509	1512	1379	889		2064	2024	1939	1782	1531	889
400	EX.....	1944	1945	1933	1713	1081		2568	2507	2381	2192	1858	1087
	CG.. .	1920	1926	1913	1698	1065		2567	2506	2388	2178	1857	1065
	CD ..	1920	1928	1918	1703	1065		2567	2509	2394	2186	1865	1065
500	EX... .	2353	2353	2321	2026	1256		3017	2950	2790	2556	2158	1259
	CG ...	2301	2302	2275	1990	1228		3015	2940	2795	2541	2158	1228
	CD.	2301	2303	2280	1998	1228		3015	2943	2802	2551	2168	1228
550	EX . .	5556?	2542	2506	2173	1341		3220	3147	2978	2733	2288	1381
	CG . .	2479	2478	2445	2130	1308		3224	3140	2986	2714	2302	1308
	CD .. .	2479	2479	2450	2137	1308		3224	3146	2993	2721	2313	1308

VISCOSITY CO_2–N_2O							VISCOSITY O_2–CO					
T	%CO_2 . . .	0.0	40.24	60.33	80.97	100.0	%O_2. .	0.0	23.37	42.01	77.33	100.0
°*K*.												
300	EX. . .	1488	1494	1495	1490	1493		1776	1841	1900	1998	2057
	CG.....	1489	1497	1498	1496	1493		1779	1880	1946	2034	2064
	CD .. .	1489	1523	1524	1513	1493		1779	1844	1897	1998	2064
400	EX.....	1493?	1950	1950	1941	1944		2183	2268	2343	2482	2568
	CG.....	1936	1938	1934	1928	1920		2210	2328	2407	2521	2567
	CD ...	1936	1965	1961	1946	1920		2210	2292	2358	2485	2567
500	EX.....	2355	2365	2365	2358	2353		2548	2650	2741	2908	3017
	CG.....	2338	2332	2325	2314	2301		2593	2725	2817	2954	3015
	CD.....	2338	2361	2354	2332	2301		2593	2690	2768	2918	3015
550	EX.....	2555	2562	2564	2551	2556						
	CG.....	2525	2517	2508	2495	2479						
	CD.....	2525	2546	2537	2513	2479						

VISCOSITY C_2H_4–N_2							VISCOSITY H_2–N_2O					
T	%C_2H_4 ...	0.0	24.05	56.95	76.21	00.0	%H_2....	0.0	39.89	59.61	78.57	100.0
°*K*.												
300	EX.....	1781	1574	1308	1169	1033		1488	1481	1451	1348	891
	CG...	1777	1562	1306	1176	1029		1489	1509	1483	1376	889
400	EX... .	2190	1956	1655	1491	1348		1943	1907	1849	1684	1081
	CG... .	2192	1946	1653	1501	1330		1936	1932	1876	1710	1065
500	EX.....	2560	2292	1963	1786	1622		2355	2292	2206	1990	1256
	CG.....	2556	2284	1961	1792	1602		2338	2311	2229	2009	1228
550	EX....	2727	2453	2108	1921	1753		2555	2477	2376	2137	1341
	CG.....	2729	2444	2104	1927	1728		2525	2489	2392	2150	1308

* EX = experimental data; CG = calculated using a geometric mean ϵ/k; CD = calculated using an ϵ/k calculated from diffusion.

TABLE 6—*Continued*

	VISCOSITY He-A					VISCOSITY H_2-CH_4						
T	%He . . .	0.0	38.20	49.06	100.0	%H_2 . .	0.0	28 08	48.55	60.22	92.23	100.0
°K.												
293	EX*. . .	2211	2291	2296	1973		1087	1099	1098	1086	955	876
	CG*. . . .	2223	2305	2319	1926		1092	1102	1096	1083	952	876
	CD*. . . .	2223	2314	2330	1926		1092	1105	1101	1088	955	876
373	EX. . . .	2684	2745	2750	2320		1331	1337	1328	1306	1132	1033
	CG.	2688	2753	2757	2244		1327	1329	1313	1291	1115	1021
	CD.	2688	2765	2774	2244		1327	1333	1318	1297	1119	1021
473	EX.	3208	3250		2715		1603	1602	1587	1551	1338	1213
	CG.	3205	3261	3257	2615		1588	1583	1557	1525	1300	1185
	CD.	3205	3282	3285	2615		1588	1585	1561	1531	1303	1185
523	EX.	3448	3488		2903		1725	1718	1699	1662	1423	1296
	CG.	3443	3502	3496	2790		1709	1701	1670	1634	1389	1265
	CD.	3443	3518	3516	2790		1709	1704	1675	1641	1393	1265

	VISCOSITY A-Ne						VISCOSITY He-Ne					
T	%A	0.0	26.80	60.91	74.20	100.0	%He . .	0.0	26.59	56.24	76.21	100.0
°K.												
293	EX.	3092	2808	2504	2401	2213		3092	2971	2702	2429	1941
	CG.	3079	2805	2503	2401	2223		3079	2935	2675	2407	1926
373	EX.	3623	3313	2990	2885	2693		3623	3479	3171	2846	2281
	CG.	3602	3306	2988	2878	2688		3602	3428	3122	2805	2244
473	EX.	4220	3890	3529	3413	3222		4220	4056	3702	3327	2672
	CG.	4168	3862	3526	3410	3205		4168	3980	3633	3271	2615
523	EX. . . .	4501	4150	3793	3658	3460		4501	4310		3555	2853
	CG.	4447	4127	3779	3656	3443		4447	4244	3875	3488	2790

	VISCOSITY H_2-Ne						VISCOSITY H_2-A					
T	%H_2. . . .	0.0	25.20	46.09	77.15	100.0	%H_2 . .	0.0	29.42	44.57	65.15	100.0
°K.												
293	EX.	3092	2782	2427	1684	875		2211	2140	2056	1857	875
	CG.	3079	2763	2411	1670	876		2223	2140	2060	1860	876
373	EX.	3623	3269	2845	1981	1029		2684	2586	2488	2238	1029
	CG.	3602	3231	2820	1950	1021		2688	2576	2470	2219	1021
473	EX.	4220	3807	3327	2319	1211		3208	3070	2948	2636	1211
	CG.	4168	3741	3267	2263	1185		3205	3063	2929	2619	1185
523	EX.	4501	4054	3540	2476	1296		3448	3310	3164	2826	1296
	CG.	4447	3992	3487	2416	1265		3443	3287	3143	2808	1265

* EX = experimental data; CG = calculated using a geometric mean ϵ/k; CD = calculated using an ϵ/k calculated from diffusion.

TABLE 6—*Continued*

	VISCOSITY N_2–O_2						VISCOSITY CO–N_2						
T	%N_2	0.0	24.08	58.93	78.22	100.0	%CO.	0.0	18.46	39.70	65.68	83.71	100.0
°K.													
300	EX*....	2057	1995	1894	1843	1781		1781	1782	1781	1775	1774	1776
	CG*....	2064	1994	1894	1839	1777		1779	1826	1855	1852	1823	1777
	CD*...	2064	2015	1921	1858	1777		1779	1778	1780	1780	1780	1777
400	EX.....	2568	2480	2345	2275	2190		2190	2186	2183	2191	2184	2183
	CG.....	2568	2475	2343	2272	2192		2210	2245	2279	2279	2252	2192
	CD .. .	2568	2497	2372	2292	2192		2210	2196	2201	2205	2208	2192
500	EX....	3017	2909	2741	2658	2560		2560	2560	2558	2549	2551	2548
	CG.....	3015	2901	2741	2654	2556		2593	2617	2656	2661	2636	2556
	CD .. .	3015	2923	2768	2674	2556		2593	2564	2573	2582	2588	2556
550	EX.....	3220	3109	2932	2840	2727		2727	2721	2719	2722	2719	2714
	CG.....	3224	3100	2925	2833	2729		2771	2792	2834	2840	2814	2729
	CD	3224	3125	2956	2854	2729		2771	2737	2746	2757	2763	2729

	VISCOSITY C_2H_4–CO						VISCOSITY H_2–C_3H_8				
T	%C_2H_4....	0 0	26.32	43.54	80.62	100.0	%H_2 .	0.0	37.04	78.82	92.25	100.0
°K.											
300	EX.....	1776	1553	1402	1135	1033		817	874	985	970	891
	CG... .	1779	1538	1399	1144	1029		819	892	995	980	889
400	EX.....	2183	1943	1763	1460	1342		1070	1130	1233	1194	1081
	CG	2210	1930	1768	1466	1330		1074	1152	1241	1194	1065
500	EX.....	2548	2279	2098	1760	1622		1308	1366	1459	1392	1256
	CG. ...	2593	2279	2097	1757	1602		1309	1390	1462	1389	1228
550	EX.. ..	2714	2433	2240	1900	1753		1422	1478	1566	1485	1347
	CG.....	2771	2440	2249	1891	1728		1418	1499	1566	1482	1308

	VISCOSITY CO_2–C_3H_8						VISCOSITY N_2O–C_3H_8				
T	%CO_2 . .	0 0	40.25	57.76	78.83	100.0	%N_2O..	0.0	20.16	58.29	79.82	100.0
°K.											
300	EX.....	817	1058	1174	1326	1493		817	926	1167	1326	1488
	CG.....	819	1033	1146	1304	1493		819	917	1145	1307	1489
400	EX.....	1070	1383	1533	1730	1944		1070	1213	1525	1725	1943
	CG.....	1074	1345	1487	1685	1920		1074	1201	1494	1703	1936
500	EX.....	1308	1670	1856	2093	2353		1308	1478	1854	2083	2355
	CG.. ..	1309	1627	1795	2026	2301		1309	1460	1811	2059	2338
550	EX.....	1422	1815	2010	2267	2556		1422	1610	2012	2271	2556
	CG.....	1418	1759	1939	2186	2479		1418	1581	1959	2227	2525

* EX = experimental data; CG = calculated using a geometric mean ϵ/k; CD = calculated using an ϵ/k calculated from diffusion.

TABLE 6—*Concluded*

VISCOSITY H_2–N_2

T	%H_2.	0.0	25.00	50.00	75 00	100.0
°K.						
82	EX*...	544	540	524	493	362
	CG*....	564	564	552	508	356
292	EX.....	1746	1700	1609	1396	882
	CG.....	1744	1699	1605	1393	874

VISCOSITY C_2H_4–O_2

T	%C_2H_4 .	0.0	13.06	41.45	77.03	100.0
°K.						
293	EX.....	2019	1854	1529	1198	1010
	CG....	2027	1850	1521	1187	1007
323	EX.....	2181	2004	1658	1308	1107
	CG....	2187	2001	1650	1294	1102
373	EX.....	2433	2243	1865	1479	1262
	CG.. .	2440	2238	1855	1465	1254

VISCOSITY SO_2–CO_2

T	%SO_2	0.0	20.00	40.00	60.00	80.00	100.0
°K.							
289	EX .	1458	1428	1388	1346	1299	1243
	CG....	1443	1420	1397	1375	1353	1331

VISCOSITY CO_2–HCl

T	%CO_2. .	0.0	20.00	40.00	60.00	80.00	100.0
°K.							
291	EX....	1426	1453	1473	1483	1481	1464
	CG....	1422	1429	1435	1442	1448	1453

VISCOSITY air–CO_2

T	%Air.....	0.0	20.00	40.00	60.00	80.00	100.0
°K.							
290	EX....	1455	1523	1591	1660	1730	1797
	CG....	1448	1511	1578	1648	1722	1800

VISCOSITY air–HCl

T	%Air.	0.0	20.00	40.00	60.00	80.00	100.0
°K.							
290	EX....	1407	1535	1616	1693	1755	1794
	CG....	1415	1499	1582	1660	1734	1798

* EX = experimental data; CG = calculated using a geometric mean ϵ/k; CD = calculated using an ϵ/k calculated from diffusion.

$$J_{ii} = +A_{ii} - B_{ii} - \sum^{f} (x_k/x_i)(A_{ik} + B_{ik}M_k/M_i) \qquad (23)$$

and the A_{ij} and B_{ij} are defined as:

$$A_{ij} = (0.017663)T^{-1/2}[r_{ij}]^2[W^{(1)}(1;\, kT/\epsilon_{ij})][M_iM_j/(M_i + M_j)^3]^{1/2} \qquad (24)$$

$$B_{ij} = (0.0052988)T^{-1/2}[r_{ij}]^2[W^{(2)}(2;\, kT/\epsilon_{ij})][M_iM_j/(M_i + M_j)^3]^{1/2} \qquad (25)$$

In table 7a are shown some experimental figures for the viscosity of a helium–neon–argon mixture; along with this data are given the calculated results obtained by using equation 21, letting $f = 3$. The force constants between like molecules which were used are those obtained from viscosity data; the force constants between unlike molecules which were used are those obtained by taking an arithmetic-mean collision diameter and a geometric-mean ϵ/k.

The methods presented here are well suited to the prediction of transport properties for such industrially important mixtures as flue and fuel gases. In table 7b we compare the experimentally observed values of the viscosity of flue and fuel gases with values calculated using equation 21, together with the force constants given in tables 1 and 2. The agreement in more than half of the examples is excellent; however, in some of the other mixtures there is a considerable discrepancy. This might be explained on the basis of small errors in the experimental gas analysis. In any case we believe that our calculated values have an inherent accuracy of within 0.5 per cent. The fact that some of these mixtures contained as many as seven components caused no computational difficulties.

In this report we do not consider quantum corrections, inasmuch as they are numerically small and negligible at any reasonable temperature ($T > 15$°K.).

TABLE 7A

Viscosity of ternary mixture of neon–argon–helium

T	VOLUME PER CENT			$\eta_{123} \times 10^7$ g./cm.-sec.	
	Ne	A	He	Calculated	Experimental(a)
°K.					
193	55.76	26.70	17.54	2718	2740
	31.93	32.13	35.94	2562	2569
	21.66	58.51	19.83	2429	2411
	21.89	23.82	54.29	2500	2504
373	55.76	26.70	17.54	3205	3237
	31.93	32.13	35.94	3025	3044
	21.66	58.51	19.83	2895	2886
	21.89	23.82	54.29	2938	2957
473	55.76	26.70	17.54	3752	3790
	31.93	32.13	35.94	3551	3574
	21.66	58.51	19.83	3425	3415
	21.89	23.82	54.29	3449	3470

Force constants used

LIKE MOLECULES			UNLIKE MOLECULES		
Molecule	r_i	ϵ_i/k	Molecule pair	Arithmetic mean r_{ij}	Geometric mean ϵ_{ij}/k
	Å.	°K.			
Ne	2.80	35.7	Ne–A	3.11	68
A	3.418	124.0	Ne–He	2.75	14
He	2.70	6.03	A–He	3.059	27

(a) M. Trautz and K. F. Kipphan: Ann. Physik **2**, 746 (1929).

The quantum corrections would be expected to be larger for hydrogen than for any heavier molecules. The magnitude of this correction may be seen from a comparison of the viscosity of H_2 and D_2. If there were no quantum corrections, η_{H_2} would equal $(M_H/M_D)^{1/2}\eta_{D_2}$. This would be true regardless of the form for the energy of interaction as long as classical mechanics applies, but quantum corrections would introduce deviations. From table 8 it is seen that there is no appreciable deviation (nor quantum correction) down to 15°C. Similarly in the case of CH_4 and CD_4, where it is expected that $\eta_{CH_4} = (M_{CH_4}/M_{CD_4})^{1/2}\eta_{CD_4}$. Here,

TABLE 7B

Viscosity of multicomponent mixtures: industrial gases

COMPOSITION (VOLUME PER CENT)							TEMPERATURE	$\eta \times 10^7$ EXPERIMENTAL	$\eta \times 10^7$ CALCULATED	REFERENCE
CO_2	O_2	CO	H_2	CH_4	N_2	Heavier hydrocarbons				
							°K.			
8.6	2.3				89.1		293	1756	1761	(b)
13.3	3.9				82.8		293	1749	1765	(b)
6.2	10.7				83.1		293	1793	1789	(b)
10.4		28.5	1.6		59.5		293	1738	1798	(b)
10.80	2.00		2.20		85.00		300.5	1827	1792	(a)
							524.5	2715	2661	
							973	4117	4008	
							1279	4856	4753	
6.70	0.10	7.80	2.20		83.20		307.5	1842	1835	(a)
							519	2655	2653	
							975	4048	4019	
							1285	4808	4783	
6.40	3.00	0.30	0.70		89.60		314	1904	1856	(a)
							518	2706	2644	
							974.5	4113	4017	
							1287	4895	4777	
6.00	0.10	25.70	11.50		56.70		302	1823	1829	(a)
							526	2686	2696	
							976	4041	4042	
							1283	4777	4821	
10.6		29.8	3.9	0.3	55.4		293	1743	1794	(b)
8.9		30.7	3.3	0.4	56.7		293	1747	1797	(b)
8.7		32.8	1.5	0.2	56.8		293	1749	1802	(b)
3.70	0.30	27.10	9.50	1.60	57.80		300.5	1815	1816	(a)
							565.5	2819	2823	
							981	4045	4041	
							1282	4792	4803	
1.7	0.9	6.0	57.5	24.0	7.8	2.1	293	1262	1254	(b)
2.1	0.9	5.7	53.0	24.3	11.7	2.3	293	1304	1290	(b)
2.0	1.4	4.6	54.9	23.5	11.6	2.0	293	1310	1398	(b)
3.3	0.6	3.8	51.3	29.6	10.0	1.4	293	1332	1269	(b)
2.2	0.6	4.1	53.1	29.5	9.2	1.3	293	1306	1254	(b)
2.2	1.0	4.0	52.3	29.9	9.4	1.2	293	1307	1261	(b)
2.5	0.8	14.9	53.0	18.1	9.1	1.6	293	1355	1373	(b)
4.8	0.3	26.4	17.2	2.6	48.2	0.5	293	1714	1743	(b)
3.5	0.3	27.3	14.4	3.7	50.0	0.8	293	1712	1732	(b)
3.1	0.5	28.6	17.7	4.2	45.0	0.9	293	1715	1719	(b)

(a) Schmid: Gas- u. Wasserfach **85**, 92 (1942).
(b) F. Herning and L. Zipperer: Gas- u. Wasserfach **79**, 49–54, 69–73 (1936).

from the figures shown in table 8, there is no appreciable quantum correction indicated by the experimental data down to 90°K. Only in the cases of hydrogen and helium is it expected that the transport properties would show a small quantum correction above 15°K. In these two cases it appears probable that a part of our difficulty in obtaining a suitable fit with the experimental viscosity data may arise from small quantum corrections. The conclusions have been substantiated by the theoretical work of Uehling and Uhlenbeck (4, 16, 17) and Massey and Mohr (13).

TABLE 8

Viscosity of light and heavy methane and light and heavy hydrogen

T	$\sqrt{\frac{M_{H_2}}{M_{D_2}}}\,\eta_{D_2}$ (EXPERIMENTAL)	η_{H_2} (EXPERIMENTAL)(a)	η_{H_2} (CALCULATED)
°K.			
15	71	82	70
30	146	158	141
75	332	347	331
290	886	884	876

T	$\sqrt{\frac{M_{CH_4}}{M_{CD_4}}}\,\eta_{CD_4}$ (EXPERIMENTAL)	η_{CH_4} (EXPERIMENTAL)(b)	η_{CH_4} (CALCULATED)
°K.			
90	384	376	353
200	828	785	780
225	922	874	868
250	1013	961	955
275	1101	1039	1037
300	1095	1111	1115

(a) A. Van Itterbeek and O. Van Paemel: Physica 7, 265 (1940).
(b) A. Van Itterbeek: Physica 9, 831 (1940).

IV. THE COEFFICIENT OF DIFFUSION

The first approximation for the coefficient of diffusion is given by the following equation:

$$(D_{12})_1 = \frac{0.000929167T^{3/2}[(M_1 + M_2)/M_1M_2]^{1/2}}{p(r_{12})^2 W^{(1)}(1; kT/\epsilon_{12})} \tag{26}$$

in which D_{12} is the coefficient of diffusion in cm.^{2} sec.$^{-1}$, p is the pressure in atmospheres, and the M_i represent molecular weights. Note that the first approximation to the diffusion coefficient is independent of composition. This is no longer true in the case of the second approximation:

$$(D_{12})_2 = (D_{12})_1/(1 - \Delta) \tag{27}$$

Here Δ, a small quantity usually less than 0.03, is defined by

$$\Delta = 5(C - 1)^2(P_1x_1^2 + P_2x_2^2 + P_{12}x_1x_2)/(Q_1x_1^2 + Q_2x_2^2 + Q_{12}x_1x_2) \tag{28}$$

in which

$$P_1 = M_1^3[M_1 + M_2]^{-2}(2/3V_0)([D_{12}]_1/[\eta_1]_1)(273.16P/T) \tag{29}$$

$$P_2 = M_2^3[M_1 + M_2]^{-2}(2/3V_0)([D_{12}]_1/[\eta_2]_1)(273.16P/T) \tag{30}$$

$$P_{12} = [3(M_1 - M_2)^2 + 4M_1M_2A]/(M_1 + M_2)^2 \tag{31}$$

$$Q_1 = (P_1/M_1^2)(6M_2^2 + 5M_1^2 - 4M_1^2B + 8M_1M_2A) \tag{32}$$

$$Q_2 = (P_2/M_2^2)(6M_1^2 + 5M_2^2 - 4M_2^2B + 8M_1M_2A) \tag{33}$$

$$Q_{12} = \begin{bmatrix} 3(5 - 4B)(M_1 - M_2)^2/(M_1 + M_2)^2 \\ + 4AM_1M_2(11 - 4B)/(M_1 + M_2)^2 \\ + 2P_1P_2(M_1 + M_2)^4/M_1^2M_2^2 \end{bmatrix} \tag{34}$$

Here the x_i are mole fractions of the i^{th} component; the M_i are molecular weights; A, B, and C are functions of kT/ϵ_{12} tabulated previously; V_0 is the molar volume under standard conditions, i.e., 22,414.6 cc.; and $[\eta_1]_1$ and $[\eta_2]_1$ are calculated first approximations to the viscosity of the pure component given by the formula:

$$[\eta_i]_1 = 2.6693(10^{-5})(M_iT)^{1/2}/[(r_i)^2W^{(2)}(2; kT/\epsilon_i)] \tag{6'}$$

The possibility of getting good force constants between pairs of dissimilar molecules from equation 26 has already been mentioned. This equation could be written for two temperatures, inserting two diffusion measurements, and the resulting equations solved simultaneously for r_{12} and ϵ_{12}/k.

In the case of self-diffusion

$$M_1 = M_2 \qquad \epsilon_1 = \epsilon_2 = \epsilon_{12} \qquad r_1 = r_2 = r_{12}$$

and equations 26 and 27 reduce to the particularly simple forms for the first approximation to the coefficient of self-diffusion:

$$(D_{11})_1 = 0.0013140T^{3/2}/[pr_0^2W^{(1)}(1; kT/\epsilon)M^{1/2}] \tag{35}$$

and the second approximation is given by:

$$(D_{11})_2 = (D_{11})_1/(1 - \Delta) \tag{36}$$

in which Δ is defined simply as:

$$\Delta = 5(C - 1)^2/(11 - 4B + 8A) \tag{37}$$

and, as before, is a small correction factor.

The coefficients of self-diffusion may be calculated from our previously assigned values of the force constants without making any adjustments. The results are shown in table 9, along with experimental values. Harteck and Schmidt determined the self-diffusion coefficient for hydrogen by using ortho- and parahydrogen. The self-diffusion coefficients for the other gases shown were determined using gases containing tracer isotopes. In all of these cases the agreement is very good between the experimental and calculated quantities. Additional experimental data would be very useful for purposes of comparison.

The available values for the coefficient of ordinary diffusion have been analyzed to give the force constants of table 2 for the interaction between dissimilar molecules.

In a mixture containing f components the concept of diffusion may be generalized (8). If U_i is the average velocity of molecules of the i^{th} species relative to U, the mass average velocity of the gaseous mixture, then rigorously:

$$U_i = (n^2/\rho n_i) \sum_{k=1} m_k\, d_{ik} \operatorname{grad} x_j \tag{38}$$

where

$$n = \sum_{k=1}^{f} n_k = p/(RT) \tag{39}$$

s the total number of moles of gas in a cubic centimeter. The d_{ij} are not the usual diffusion constants for binary mixtures, D_{ij}, but the two are very closely related, in the general case by:

$$d_{12} = -\frac{\rho n_1}{m_2}\left[\frac{\begin{vmatrix} 0 & A_{23} & A_{24}\cdots \\ A_{32} & 0 & A_{34}\cdots \\ A_{42} & A_{43} & 0\ \cdots \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{vmatrix} + \begin{vmatrix} A_{12} & A_{13} & A_{14}\cdots \\ A_{32} & 0 & A_{34}\cdots \\ A_{42} & A_{43} & 0\ \cdots \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{vmatrix}}{\begin{vmatrix} 0 & A_{12} & A_{13} & A_{14}\cdots \\ A_{21} & 0 & A_{23} & A_{24}\cdots \\ A_{31} & A_{32} & 0 & A_{34}\cdots \\ A_{41} & A_{42} & A_{43} & 0\ \cdots \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{vmatrix}}\right] \tag{40}$$

Here the coefficients in the determinant are

$$A_{ik} = (n_i n_k/D_{ik}) + (n_k m_k/m_i) \sum_{s \neq i}^{f} (n_s/D_{is}) \tag{41}$$

The other d_{ij} may be obtained from the final form of the d_{12} by permuting the indices or from equation 40 by suitable modifications of the determinants in the numerator. This generalized formula may be obtained also from the work of Hellund and Uehling (5).

For the binary mixture

$$d_{12} = D_{12} \tag{42}$$

and for ternary mixtures:

$$d_{12} = D_{12} \quad 1 + \frac{n_3\left(\dfrac{m_3}{m_2} D_{13} - D_{12}\right)}{n_1 D_{23} + n_2 D_{13} + n_3 D_{12}}\Bigr] \tag{43}$$

TABLE 9

The coefficient of self-diffusion

GAS	T	P	D_{11}		REFERENCE FOR EXPERIMENTAL VALUES
			Calculated	Experimental	
	°K.	*atm.*	*cm.²/sec.*	*cm.²/sec.*	
Hydrogen	273	1	1.243	$1.28_5 \pm 0.002_5$	(a)
	85	1	0.167	0.172 ± 0.008	(a)
	20.4	1	0.01043	0.00816 ± 0.0002	(a)
Krypton	293	1	0.093	$0.093 \pm 0.004_5$	(b)
Xenon	293	1	0.055	$0.044 \pm 0.002_2$	(b)
Neon	293	1	0.491	0.473 ± 0.002	(c)
Argon	295	0.4211	0.422	0.423 ± 0.003	(d)
Nitrogen	293	1	0.198	0.200 ± 0.008	(e)
Methane	292	0.00811	26.38	26.32 ± 0.73	(f)
Hydrogen chloride	295	1	0.127	0.1246	(g)

[a] P. Harteck and H. W. Schmidt: Z. physik. Chem. **21B**, 447 (1933).

[b] W. Groth and P. Harteck: Z. Elektrochem. **47**, 167 (1941).

[c] W. Groth and E. Sussner: Z. physik. Chem. **193**, 296 (1944).

[d] F. Hutchinson: Phys. Rev. **72**, 1256 (1947).

[e] E. B. Winn: Phys. Rev. **74**, 698 (1948).

[f] E. B. Winn and E. P. Ney: Phys. Rev. **72**, 77 (1947).

[g] H. Braune and F. Zehle: Z. physik. Chem. **49B**, 247 (1941).

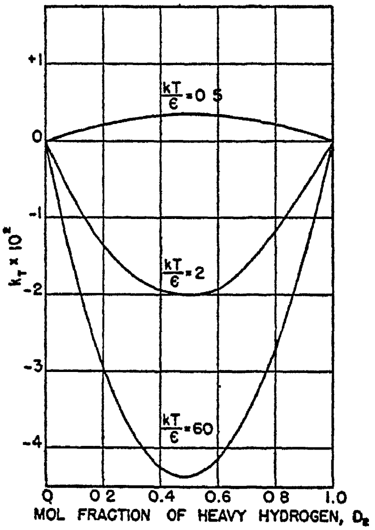


FIG. 2. The calculated thermal diffusion ratio for separation of a mixture of hydrogen and deuterium. Note the change in sign of k_T at low temperatures, an effect which has been experimentally observed for some substances.

V. THE THERMAL DIFFUSION RATIO

Since the second approximation to the thermal diffusion ratio involves very complicated algebra, only the first approximation will be set forth here. It is:

$$k_T = 5x_1x_2(C - 1)(S_1x_1 - S_2x_2)/(Q_1x_1^2 + Q_2x_2^2 + Q_3x_1x_2) \qquad (44)$$

in which

$$S_1 = \frac{M_1}{5}\left(\frac{r_1}{r_{12}}\right)^2\left[\frac{2(M_1+M_2)^3}{M_2}\right]^{1/2}\left[\frac{W^{(2)}(2; kT/\epsilon_1)}{W^{(1)}(1; kT/\epsilon_{12})}\right] - 3M_2(M_2 - M_1) - 4AM_1M_2 \quad (45)$$

$$S_2 = \frac{M_2}{5}\left(\frac{r_2}{r_{12}}\right)^2\left[\frac{2(M_1+M_2)^3}{M_1}\right]^{1/2}\left[\frac{W^{(2)}(2; kT/\epsilon_2)}{W^{(1)}(1; kT/\epsilon_{12})}\right] - 3M_1(M_1 - M_2) - 4AM_1M_2 \quad (46)$$

$$Q_1 = \frac{1}{5}\left(\frac{r_1}{r_{12}}\right)^2\left[\frac{2(M_1+M_2)}{M_2}\right]^{1/2}\left[\frac{W^{(2)}(2; kT/\epsilon_1)}{W^{(1)}(1; kT/\epsilon_{12})}\right] \cdot [6M_2^2 + (5 - 4B)M_1^2 + 8AM_1M_2] \quad (47)$$

$$Q_2 = \frac{1}{5}\left(\frac{r_2}{r_{12}}\right)^2\left[\frac{2(M_1+M_2)}{M_1}\right]^{1/2}\left[\frac{W^{(2)}(2; kT/\epsilon_2)}{W^{(1)}(1; kT/\epsilon_{12})}\right] \cdot [6M_1^2 + (5 - 4B)M_2^2 + 8AM_1M_2] \quad (48)$$

$$Q_3 = \left[3(M_1 - M_2)^2(5 - 4B) + 4AM_1M_2(11 - 4B) + \frac{4}{25}\left(\frac{r_1r_2}{r_{12}}\right)^2 \cdot \left[\frac{(M_1+M_2)^3}{(M_1M_2)^{1/2}}\right]\left[\frac{W^{(2)}(2; kT/\epsilon_1)W^{(2)}(2; kT/\epsilon_2)}{W^{(1)}(1; kT/\epsilon_{12})}\right]\right] \quad (49)$$

Here the M_i are the molecular weights of the i^{th} species; the x_i are mole fractions; the A, B, and C are functions of kT/ϵ_{12} previously tabulated; and the r_1, r_2, and r_{12} are low-velocity collision diameters in Angströms. When the thermal diffusion of isotopes is under consideration, equation 44 may be used where the S's and the Q's have the more simple form:

$$S_1 = AM_1[(2/M_2)^{1/2} - 4M_2] - 3M_2(M_2 - M_1) \quad (50)$$

$$S_2 = AM_2[(2/M_1)^{1/2} - 4M_1] - 3M_1(M_1 - M_2) \quad (51)$$

$$Q_1 = A(2/M_2)^{1/2}[6M_2^2 + (5 - 4B)M_1^2 + 8AM_1M_2] \quad (52)$$

$$Q_2 = A(2/M_1)^{1/2}[6M_1^2 + (5 - 4B)M_2^2 + 8AM_1M_2] \quad (53)$$

$$Q_3 = 4A^2(M_1M_2)^{-1/2} + 3(M_1 - M_2)^2(5 - 4B) + 4AM_1M_2(11 - 4B) \quad (54)$$

A graph of k_T *versus* composition for light and heavy hydrogen calculated with the above formula is shown in figure 2. The force constants for a mixture of isotopes are the same as those for the pure substance. It is interesting to note that these theoretical values predict the change in sign in the thermal diffusion coefficient which have been found to occur at low temperatures. At very low temperatures the thermal diffusion is so small that no practical use can be made of this inversion.

Experimental data are generally reported as the ratio, R_T, of the thermal diffusion ratio, k_T, to the corresponding value for rigid spherical molecules. While the calculated and experimental k_T are functions of temperature for real molecules, k_T for rigid spheres is independent of temperature, and it is computed using the rigid spheres quantities:

$$A = 0.4, \quad B = 0.6, \quad C = 1.2, \quad W^{(1)}(1) = \tfrac{1}{2}, \quad \text{and } W^{(2)}(2) = 1 \qquad (55)$$

TABLE 10

Thermal diffusion of gases

GAS MIXTURE	CONCENTRATION	T	R_T (CALCULATED)	R_T (EXPERIMENTAL)	REFERENCE
		°K.			
He–Ne	53.8 % He	585.2	0.71	0.64	(a)
		233.2	0.70	0.64	
		117.2	0.68	0.58	
He–A	51.2 % He	585.2	0.73	0.66	(a)
		233.2	0.71	0.63	
		117.2	0.63	0.55	
He–Xe	53.6 % He	585.2	0.75	0.66	(a)
		233.2	0.70	0.66	
Ne–A	51.2 % Ne	585.2	0.66	0.57	(a)
		233.2	0.52	0.48	
		117.2	0.27	0.28	
Ne–Xe	54.2 % Ne	585.2	0.66	0.66	(a)
		233.2	0.44	0.46	
A–Xe	56.4 % A	585.2	0.51	0.45	(a)
		233.2	0.15	0.15	
H_2–D_2		333	0.63	0.50	(b)
Ne^{20}–Ne^{22}		357	0.62	0.60	(c)
A^{36}–A^{40}		620	0.55	0.47	(c)
$C^{12}H_4$–$C^{13}H_4$		274	0.29	0.49	(b)
$N^{14}H_3$–$N^{15}H_3$		308	0.08	0.11	(d)
CH_4–H_2		506	0.68	0.67	(e)
O_2–CO_2		735	0.61	0.63	(e)
CO–CO_2		725	0.60	0.40	(e)
H_2–He		284	0.53	0.41	(b)
H_2–CO_2		596	0.70	1.03	(e)

(a) K. E. Grew: Proc. Roy. Soc. (London) **189**, 402 (1947).
(b) B. F. Murphy: Phys. Rev. **72**, 836 (1947).
(c) L. G. Stier: Phys. Rev. **62**, 548 (1942).
(d) W. W. Watson and D. Woermley: Phys. Rev. **63**, 184 (1943).
(e) N. G. Schmahl and J. Schewe: Z. Elektrochem. **46**, 210 (1940).

Table 10 shows some values of R_T (calculated) and R_T (experimental). For those gas pairs for which diffusion data were available, ϵ_{12}/k obtained from diffusion (see table 2) were used; otherwise, it was necessary to employ the geometric mean of the ϵ/k for the single components. For all of the gas pairs, r_{12} was taken to be the arithmetic mean of the r_0 for each of the two components. Force constants for ammonia were obtained from second virial coefficient data.

Since the concentration gradient which results from thermal diffusion is opposed by ordinary diffusion, a steady state is reached after a certain time interval. Van Itterbeek, Van Paemel, and Van Lierde (19) have measured thermal diffusion in gas mixtures at low temperatures as per cent change in composition at the steady state $\Delta\gamma$, which may also be calculated from the following integral:

$$\Delta\gamma = \int_{T_1}^{T_2} k_T \, d \ln T \qquad (56)$$

This integration was performed numerically, making use of equation 44 for k_T, for the gases and temperatures in table 11, where the experimental and calculated values of $\Delta\gamma$ may be compared.

TABLE 11

Thermal diffusion in gas mixtures

$\Delta\gamma$ = change in per cent H_2

GAS MIXTURE	AVERAGE VOLUME PER CENT H_2	T_1	T_2	$\Delta\gamma$ (EXPERIMENTAL)(a)	$\Delta\gamma$ (CALCULATED)
		°K.	°K.		
H_2–Ne.	35.6	290.4	90.2	6.9	7.1
	50.9	290.4	90.2	8.2	8.2
	66.7	290.4	90.2	7.7	7.9
H_2–CO.	39.7	293.3	90.2	5.6	5.9
	51.7	293.3	90.2	6.6	6.7
	73.7	293.3	90.2	5.7	6.4
H_2–He.	32.3	291.7	90.2	3.4	3.8
	50.5	291.7	90.2	4.6	4.2
	65.4	291.7	90.2	3.7	3.7
H_2–N_2.	40.9	292.1	64.4	7.6	8.1
	50.8	292.1	64.6	8.4	9.0
	78.0	292.1	64.4	6.3	8.1
H_2–O_2.	33.8	293.6	90.2	4.0	5.7
	48.2	293.6	90.2	5.5	7.1
	73.7	293.6	90.2	5.3	6.9

(a) A. Van Itterbeek, O. Van Paemel, and J. Van Lierde: Physica **13**, 231 (1947).

VI. THE COEFFICIENT OF THERMAL CONDUCTIVITY

For a single gas the coefficient of thermal conductivity, λ, through the second approximation is given in terms of the viscosity, η, and the specific heat at constant volume, C_v, by the relation:

$$\lambda = (H\eta/\mathbf{V}M)[C_v + (9/4)R] \qquad (57)$$

Here R is the usual gas constant per mole, M is the molecular weight, and H and $\mathbf{V}$ are functions of kT/ϵ previously tabulated. The ratio $H/\mathbf{V}$ is very close to unity. The factor $[C_v + (9/4)R]$ depends upon the Eucken assumption that, during a collision, thermal equilibrium is established between the translational and internal degrees of freedom in the molecule. The experimental values of λ determined by Johnston and coworkers (10, 15) provide an opportunity to check the validity of this equation. Table 12 shows the comparison of the observed *versus* the calculated value of $(H/V) = (\lambda M/\eta)[C_v + (9/4)R]^{-1}$. The agreement is fairly good in the case of helium, where there are no internal degrees of freedom. In all other cases the discrepancies are large. This indicates that the Eucken assumption is not valid; hence equation 57 must be modified so as to take into account the difficulty of transferring energy from translation to rotation and vibration. The time lag in the specific heat is another indication of this difficulty. At very high temperatures, such as occur in combustions, the molecules behave more classically, and equation 57 should apply.

Until the discrepancies in the Eucken assumption are settled, there seems little advantage in trying to apply the complicated Chapman–Enskog relationships for the heat conductivity of gaseous mixtures. Instead, we shall employ equation 57, using the viscosity and specific heat of the mixture to estimate the heat conductivity.

TABLE 12

Thermal conductivity

GAS	H/V	100°K.	200°K.	300°K.
He	Experimental(a)	0.9933	0.9897	0.9733
	Calculated	1.0042	1.0043	1.0045
CH_4	Experimental(a)	0.9676	1.0258	1.0675
	Calculated	1.0000	1.0002	1.0009
NO	Experimental(a)		0.9504	0.9974
	Calculated		1.0004	1.0013
CO_2	Experimental(a)		0.8788	0.9479
	Calculated		1.0001	1.0003
O_2	Experimental(a)	0.9529	1.0042	1.0320
	Calculated	1.0000	1.0005	1.0015
Air	Experimental(b)	0.9499	1.0022	1.0107
	Calculated	1.0001	1.0008	1.0019

(a) H. L. Johnston and E. R. Grilly: J. Chem. Phys. **14**, 233 (1946).
(b) W. J. Taylor and H. L. Johnston: J. Chem. Phys. **14**, 219 (1946).

Unfortunately, accurate experimental measurements of heat conductivity are very difficult to obtain, so that there is no possibility of developing a more satisfactory empirical formulation at the present time.

The authors wish to express their appreciation to Dr. C. F. Curtiss for allowing them to set forth the generalized viscosity and diffusion equations in this paper, and for valuable assistance throughout the course of the work. The wide scope of the present report was made possible by the efforts of a computing group, under the direction of Ruth Shoemaker. In this connection we wish to thank Mary Nordling, Doris Jayne, Phyllis Rood, Phyllis Rhoades, Marjorie Leikvold, and Ted Mesmer.

The authors also wish to thank the Navy Bureau of Ordnance for their financial support under Contract NOrd 9958.

One of the authors, R. Byron Bird, desires to acknowledge the financial help which he has received from the Wisconsin Alumni Research Foundation.

REFERENCES

(1) Chapman, S., and Cowling, T. G.: *Mathematical Theory of Non-Uniform Gases.* Cambridge University Press, Teddington (1939).

(2) Corner, Jack: Trans. Faraday Soc. **35**, 711 (1939); **35**, 1 (1939).

(3) Curtiss, C. F., and Hirschfelder, J. O.: "The Kinetic Theory of Multicomponent Systems of Gases" (University of Wisconsin, July, 1947), CF-727; to be submitted to The Journal of Chemical Physics for publication.

(4) Hellund, E. J., and Uehling, E. A.: Phys. Rev. **54**, 479 (1938); **56**, 818 (1939).

(5) Hellund, E. J., and Uehling, E. A.: Phys. Rev. **56**, 818 (1939).

(6) Hildebrand, J. H., Wakeham, H. R. R., and Boyd, R. N.: J. Chem. Phys. **7**, 1094 (1939).

(7) Hirschfelder, J. O., Bird, R. B., and Spotz, E. L.: J. Chem. Phys. **16**, 968 (1948).

(8) Hirschfelder, J. O., and Curtiss, C. F.: "The Theory of Propagation of Flames. Part I" (University of Wisconsin), CF-957; to be submitted to the Journal of Chemical Physics for publication.

(9) Hirschfelder, J. O., McClure, F. T., and Weeks, I. F.: J. Chem. Phys. **10**, 201 (1942).

(10) Johnston, H. L., and Grilly, E. R.: J. Chem. Phys. **14**, 233 (1946).

(11) Lennard-Jones, J. E., and Devonshire, A. F.: Proc. Roy. Soc. (London) **A165**, 1 (1938); **A163**, 53 (1937).

(12) London, F.: Trans. Faraday Soc. **33**, 8 (1937).

(13) Massey, H. S. W., and Mohr, C. B. O.: Proc. Roy. Soc. (London) **A141**, 434 (1933); **A144**, 188 (1934).

(14) Stockmayer, W. H.: J. Chem. Phys. **9**, 398, 863 (1941).

(15) Taylor, W. J., and Johnston, H. L.: J. Chem. Phys. **14**, 219 (1946).

(16) Uehling, E. A.: Phys. Rev. **46**, 917 (1934).

(17) Uehling, E. A., and Uhlenbeck, G. E.: Phys. Rev. **43**, 552 (1933).

(18) Van Itterbeek, A., Van Paemel, O., and Van Lierde, J.: Physica **13**, 88 (1947).

(19) Van Itterbeek, A., Van Paemel, O., and Van Lierde, J.: Physica **13**, 231 (1947).

(20) Waldman, L.: Naturwissenschaften **32**, 223 (1944).

ON THE THERMODYNAMICS OF SOLUTIONS. V

An Equation of State. Fugacities of Gaseous Solutions[1]

OTTO REDLICH and J. N. S. KWONG

Shell Development Company, Emeryville, California

Received October 25, 1948

The calculation of fugacities from P–V–T data necessarily involves a differentiation with respect to the mole fraction. The fugacity rule of Lewis, which in effect would eliminate this differentiation, does not furnish a sufficient approximation. Algebraic representation of P–V–T data is desirable in view of the difficulty of numerical or graphical differentiation. An equation of state containing two individual coefficients is proposed which furnishes satisfactory results above the critical temperature for any pressure. The dependence of the coefficients on the composition of the gas is discussed. Relations and methods for the calculation of fugacities are derived which make full use of whatever data may be available. Abbreviated methods for moderate pressure are discussed.

The old problem of the equation of state has a practical aspect which has become increasingly important in recent times. The systematic description of gas reactions under high pressure requires information about the fugacities, derived sometimes from extensive data but frequently from nothing more than the critical pressure and temperature.

For practical purposes a representation of the relation between pressure, volume, and temperature based on two or three individual coefficients is desirable, although a representation of this type satisfies the theorem of corresponding states and therefore cannot be accurate. Since the calculation of fugacities involves a differentiation with respect to the mole fraction, an *algebraic* representation by means of an equation of state appears to be desirable. After this paper had been written, Joffe (12) showed that fugacities can be correctly derived also from generalized charts. An algebraic method, however, if based on a suitable equation of state, appears to be more convenient. It is free from the arbitrariness necessarily involved in the drawing of generalized charts. In addition, an approximate equation of state can be used for the precise representation of experimental data with the aid of approxiate deviation functions.

I. AN EQUATION OF STATE

Several theoretical and practical considerations lead to the following equation of state. We omit a detailed discussion because the reasoning is circumstantial and by no means rigorous. The equation is therefore essentially empirical. Its justification rests mainly on the degree of approximation obtained by comparatively simple means.

---

[1] Presented at the Symposium on Thermodynamics and Molecular Structure of Solutions, which was held under the auspices of the Division of Physical and Inorganic Chemistry at the 114th Meeting of the American Chemical Society, Portland, Oregon, September 13 and 14, 1948.

The proposed equation is represented by

$$P = RT/(V - b) - a/T^{1/2}V(V + b) \tag{1}$$

It will be used in the form:

$$Z = 1/(1 - h) - (A^2/B)h/(1 + h) \tag{2}$$
$$Z = PV/RT \tag{3}$$
$$A^2 = a/R^2T^{2.5} = 0.4278T_c^{2.5}/P_cT^{2.5} \tag{4}$$
$$B = b/RT = 0.0867T_c/P_cT \tag{5}$$
$$h = BP/Z = b/V \tag{6}$$

The relations between the coefficients a and b or A^2 and B and the critical temperature T_c and pressure P_c follow from the critical conditions. The quantities A^2 and B are expressed in atm.$^{-1}$

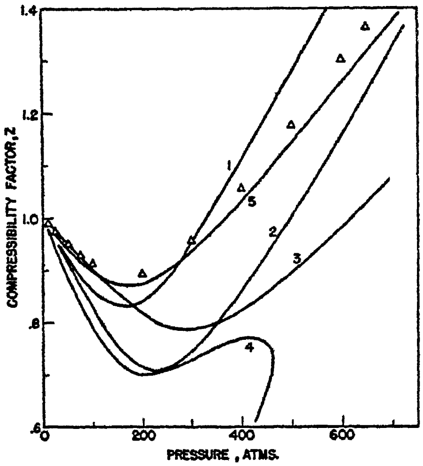


FIG. 1

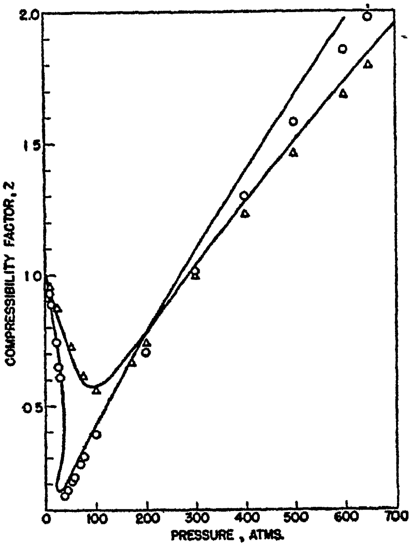


FIG. 2

FIG. 1. Compressibility factor of ethane (237.8°C.). Curve 1, van der Waals; curve 2, Dieterici; curve 3, Berthelot; curve 4, Wohl; curve 5, equation 2; △, experimental data (26).

FIG. 2. Compressibility factor of n-butane. Experimental data (23): ○, 137.8°C.; △, 237.8°C. Curves: equation 2.

The results obtained from equation 2 have been compared with experimental data for various gases in a wide pressure range. A few examples are shown in figures 1–3. The temperature, the pressure range, and the data for the maximum deviation are given in table 1 for all gases examined. A comparison with similarly simple equations of state is shown in figure 1 for the arbitrarily chosen example of ethane. Considering the limitation imposed on any equation containing but two or three individual coefficients by the failure of the correspondence theorem, the approximation obtained by equation 2 is satisfactory.

At high pressures the volume of all gases approaches a limiting value (13) which is practically independent of the temperature and close to $0.26V_c$ (V_c = critical volume). In equation 1 this limiting volume is represented by the coefficient b. The equation has been constructed to satisfy the condition

$$b = 0.26V_c \tag{7}$$

in order to furnish good approximation at high pressure.

For low pressures Berthelot's equation is known to furnish a very good approximation. The second virial coefficient derived from equation 2

$$\beta = b - a/RT^{1.5} = 0.0867(RT_c/P_c)[1 - 4.93(T_c/T)^{1.5}] \tag{8}$$

agrees quite well with the value derived from Berthelot's equation

$$\beta = 0.0703(RT_c/P_c)[1 - 6(T_c/T)^2] \tag{9}$$

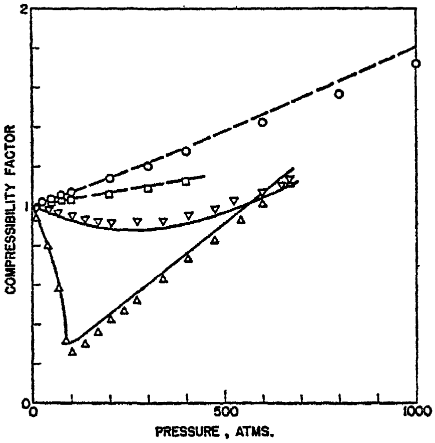


Fig. 3. Compressibility factor of hydrogen and carbon dioxide. Experimental data: hydrogen (17), ○, 0°C.; (2), □, 399.3°C. Carbon dioxide (25), ▽, 33.8°C.; △, 237.8°C. Curves: equation 2.

for the critical and higher temperatures. Below the critical temperature, the deviations of both equations from experimental data increase with decreasing temperature. Here Berthelot's equation furnishes a better approximation for low pressures.

Sometimes it is useful to know the limiting tangents in a diagram of Z against P. If we transform equation 2 into

$$Z = 1 + BP - A^2P(Z - BP)/Z(Z + BP) \tag{10}$$

we find that the tangents are represented by

$$1 + (B - A^2)P = 1 + \beta P/RT \text{ for } P = 0 \tag{11}$$

$$1 + BP \qquad \text{for } P = \infty \tag{12}$$

Although equations 2 and 6 represent Z only implicitly as a function of P, numerical calculations are not laborious: Calculate B and A^2/B according to

TABLE 1

Maximum deviations

SUBSTANCE		TEMPERATURE	PRESSURE RANGE	DATA REFERENCE	MAXIMUM DEVIATION		
					At pressure	Z (observed)	Z (calculated)
		°C.	atm.		atm.		
Hydrogen		0	2542	(17)	2542	2.706	3.067
		150	2970	(17)	2970	2.267	2.564
		399.3	400	(2)	300	1.089	1.098
Nitrogen		0	1000	(2)	1000	2.064	1.977
		399.3	1000	(2)	1000	1.511	1.416
Oxygen		0	1000	(11)	1000	1.736	1.664
		199.5	900	(11)	900	1.413	1.362
Carbon dioxide		37.8	680	(25)	680.5	1.116	1.168
		237.8	680	(25)	680.5	0.916	0.878
Methane		37.8	680	(22)	680.5	1.447	1.415
		237.8	680	(22)	680.5	1.330	1.289
Ethane		37.8	680	(26)	61.24	0.248	0.292
		237.8	680	(26)	680.5	1.405	1.353
Ethylene		150	3000	(15, 18)	2473	3.653	3.584
Propane		104.4	204	(27)	40.84	0.525	0.584
		100	181	(5)	45	0.245	0.430*
		275	303	(5)	130	0.808	0.787
n-Butane		137.8	680	(23)	680.5	2.060	2.187
		237.8	680	(23)	680.5	1.858	1.913
Isobutane		137.8	340	(19)	34.02	0.216	0.269
		237.8	340	(19)	340.2	1.117	1.094
MIXTURES	MOLE FRACTION METHANE						
Methane–ethane	0.4006	50	60	(16)	60	0.759	0.780
Methane–isobutane	0.4681	137.8	340	(21)	170.1	0.690	0.714
Methane–carbon dioxide	0.4055	37.8	680	(25)	51.0	0.641	0.613

* The value of Z changes rapidly in the vicinity of the critical point. The experimental value $Z = 0.430$ corresponds to a pressure of 43 atm.

TABLE 2

Auxiliary functions

h	$1/(1-h)$	$h/(1+h)$	h	$1/(1-h)$	$h/(1+h)$
0.02	1.0204	0.01961	0.3	1.4286	0.23077
0.04	1.0417	0.03846	0.4	1.6667	0.28571
0.06	1.0638	0.05660	0.5	2.0000	0.33333
0.08	1.0870	0.07407	0.6	2.5000	0.37500
0.1	1.1111	0.09091	0.7	3.3333	0.41176
0.15	1.1765	0.13043	0.8	5.0000	0.44444
0.2	1.2500	0.16667	0.9	10.0000	0.47369
0.25	1.3333	0.20000			

equations 4 and 5, estimate the required range of the auxiliary variable h from equation 6, calculate Z according to equation 2 with the aid of table 2 for a suitable set of values of h, and calculate $P = hZ/B$ for the values of h chosen in the preceding step. Graphical interpolation furnishes Z for any value of P.

Equation 2 can be advantageously used for the presentation of experimental data and for subsequent thermodynamic calculations. The deviation function

$$W = Z_E - Z_S \tag{13}$$

(Z_E = experimental value of Z; Z_S = value calculated from the equation of state) is restricted to a small range and varies only slowly with P. It can be easily interpolated and represented graphically or by means of a short table. Thermodynamic calculations can be carried out algebraically in Z_S so that graphical or numerical differentiations and integrations are required only for the small deviation W. The advantage is considerable, especially in differentiations.

It is convenient to call the quotient of the fugacity and the pressure "fugacity coefficient."[2] The relationship between fugacity, volume, and pressure (*cf.* 9) together with equation 2 furnishes for the fugacity coefficient φ the equation:

$$\ln \varphi = \int_0^P (Z - 1)\mathrm{d}P/P = Z - 1 - \ln (Z - BP) - (A^2/B) \ln (1 + BP/Z) \tag{14}$$

II. GASEOUS MIXTURES

It will be expected that equation 2 holds about equally well for gaseous mixtures. The question is only in which way the coefficients A and B of the mixture depend on the corresponding coefficients A_i and B_i and the mole fractions y_i of the components.

Since the coefficient b has been introduced as the limiting volume, it is quite obvious that it depends linearly on the mole fractions. Therefore,

$$b = \sum_i y_i b_i; \qquad B = \sum_i y_i B_i \tag{15}$$

This relation deviates from the original theory of van der Waals and from the general theory of the second virial coefficient (8). But the theory is valid only for moderate pressures where the influence of b is small. The linear relation 15 follows from our interpretation of b for high pressures, where the influence of b is decisive. This relation is well supported by experimental data (20).

The situation is different for the attraction coefficient a. Molecular theory leads to the result that the second virial coefficient is a function of second degree of the mole fractions. This result is also experimentally well established (8). Since the coefficient a is important at moderate pressures, a reasonable approximation can be achieved only by assuming a similar relationship for a: namely,

$$a = a_1 y_1^2 + a_2 y_2^2 + \cdots + 2a_{12} y_1 y_2 + \cdots \tag{16}$$

---

[2] Prigogine and Defay (24) recently introducted this term by means of a slightly different definition which appears to be less convenient.

The coefficients a_{12} .... of the cross terms can be related to the properties of the pure components by the usual though somewhat arbitrary assumption that

$$a_{12} = (a_1 a_2)^{1/2} \qquad (17)$$

Equations 4, 16, and 17 furnish

$$A = \sum_i y_i A_i \qquad (18)$$

Equations 15 and 18 are in accord with the conclusions of several authors, especially Gillespie (9a, 14) and Beattie (4), and are for moderate pressures well supported by their discussions of experimental data.

The second virial coefficient

$$\beta = RT(B - A^2) = RT[\Sigma y_i B_i - (\Sigma y_i A_i)^2] \qquad (19)$$

is a linear function of the mole fractions only in the exceptional case that all coefficients A_i are equal. In first approximation for low pressures we have

$$Z = 1 + \beta P/RT \qquad (20)$$

Since the attraction coefficient A^2 is in general far more important here than B, we arrive at the conclusion that the deviations from the perfect law cannot even in first approximation be represented by a linear function of the mole fractions. In general, the volume of a gaseous mixture ceases to be additive as soon as the gas ceases to be perfect or, a gaseous mixture is a perfect *solution* only as long as it is a perfect *gas*. In other words, no legitimate pressure range exists for the application of the rule of Amagat (1). Since the fugacity rule of Lewis and Randall is thermodynamically connected with Amagat's rule, this fugacity rule, too, does not furnish a correct first approximation. This conclusion does not depend on the assumption of a particular equation of state. It is in accord with experimental data (*cf.* 10).

The difference in the dependence of B and A^2 on the composition and the difference in the influence of these coefficients on Z in different pressure ranges accounts for a fact which at first is very surprising: volume–composition curves approach straight lines at high pressure but in general deviate widely from straight lines at moderate pressures.

III. THE FUGACITY COEFFICIENT

The replacement of the fugacity by the fugacity coefficient is especially useful in calculations pertaining to solutions. We define the fugacity coefficient φ_r of the component r of a solution by means of

$$\varphi_r = f_r / y_r P \qquad (21)$$

where f_r denotes the fugacity of the component r.

For gaseous mixtures the fugacity coefficient is given, therefore, by the relation

$$\ln \varphi_r = \int_0^P (\bar{Z}_r - 1) \mathrm{d}P/P \qquad (22)$$

where the partial molal compressibility factor $\bar{Z}_r$ is to be derived from Z according to

$$\bar{Z}_r = Z + \partial Z/\partial y_r - \sum_i y_i \partial Z/\partial y_i \tag{23}$$

The summation is to be extended over all components. The partial differentiation is to be carried out, of course, at constant values of all mole fractions except that one with respect to which Z is differentiated.[3]

The derivation of the fugacity coefficients from equations 2, 6, 15, 18, 22, and 23 is possible but tedious. Much more convenient is a calculation in which the order of the differentiation with respect to the mole fraction (equation 23) and the integration with respect to the pressure (equation 22) is reversed. We define the molal quantity U by

$$U = \Sigma y_i \log \varphi_i \tag{24}$$

According to equations 22 and 23, and 2 and 6, we find

$$\begin{aligned} 2.303U &= \int_0^P (Z - 1)\mathrm{d}P/P \\ &= Z - 1 - \ln (Z - BP) - (A^2/B) \ln (1 + BP/Z) \qquad (25) \end{aligned}$$

$$\log \varphi_r = U + \partial U/\partial y_r - \sum_i y_i \partial U/\partial y_i \tag{26}$$

The dependence of Z on A and B is given by equations 2 and 6, and that of A and B on the mole fractions by equations 15 and 18. The calculation leads without difficulty to

$$\begin{aligned} \log \varphi_r = 0.4343(Z - 1)B_r/B - \log (Z - BP) \\ - \frac{A^2}{B}\left[\frac{2A_r}{A} - \frac{B_r}{B}\right] \log (1 + BP/Z) \qquad (27) \end{aligned}$$

Properly arranged, the computation of fugacity coefficients according to equation 27 does not require an undue amount of time. For each temperature and composition of the gaseous mixture the fugacity coefficients of all components and for all desired pressures are calculated in one set. The required steps are:

I. A_i and B_i for all components (equations 4 and 5)
II. A and B for the mixture (equations 15 and 18)
III. Z for the mixture as a function of P (equations 2 and 6)
IV. $\log \varphi_r$ (equation 27)

Sometimes a sufficient approximation, at least for first information, is obtained by calculating the case in which one of the components is present in large excess, $y_s = 1$. In this case the quantities A, B, and Z assume the values

---

[3] Equation 23 applies to any molal quantity Z and the corresponding partial molal quantities $\bar{Z}_r$. It is derived from the usual definition of partial molal quantities by replacing mole numbers by mole fractions. The derivation of equation 23 is *not* based on an assumption that the y_i's are independent. It is therefore not necessary, though of course it is permissible, to eliminate one of the y_i's in Z before applying equation 23.

A_s, B_s, and Z_s of the pure solvent gas, so that the calculation is considerably shorter.

TABLE 3

Hydrogen–nitrogen

TEMPERATURE	MOLE FRACTION H_2	0		0.5		1	
0°C.	Pressure (atm.)	600	1000	600	1000	600	1000
	Fugacity of components (calculated)	689	1636			927	2159
	Same (Bolshakov)	737	1811			899	1992
	φ_1/φ_1^0 (calculated)	1.27	1.26	1.09	1.08	1	1
	φ_1/φ_1^0 (Bolshakov)	1.20	1.20	1.07	1.07	1	1
	φ_2/φ_2^0 (calculated)	1	1	1.03	1.07	1.28	1.32
	φ_2/φ_2 (Bolshakov)	1	1	1.04	1.05	1.23	1.27
200°C.	Fugacity of components (calculated)	758	1569			783	1565
	Same (Bolshakov)	831	1806			769	1513
	φ_1/φ_1^0 (calculated)	1.08	1.09	1.03	1.04	1	1
	φ_1/φ_1^0 (Bolshakov)	1.03	1.04	1.02	1.02	1	1
	φ_2/φ_2^0 (calculated)	1	1	1.01	1.01	1.08	1.11
	φ_2/φ_2^0 (Bolshakov)	1	1	1.00	1.00	1.04	1.04

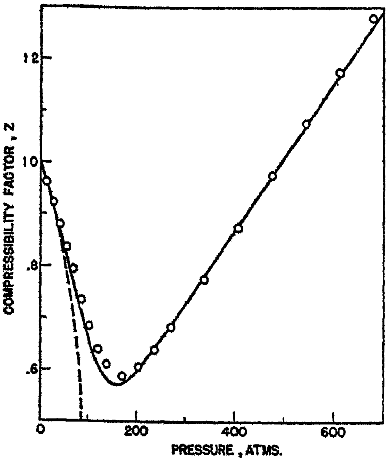


FIG. 4

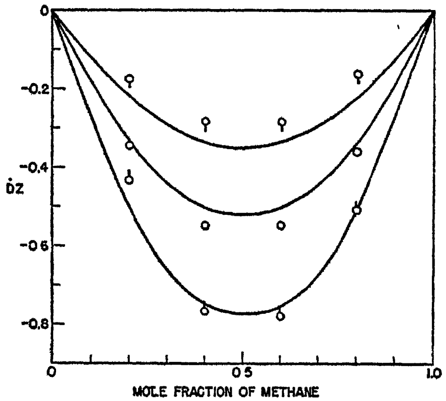


FIG. 5

FIG. 4. Compressibility factor of methane–carbon dioxide (mole fraction of methane = 0.4055). Experimental data (25): O, 37.8°C. Full line: equations 2, 15, 18. Broken line: Berthelot's equation.

FIG. 5. Methane–ethane (50°C.). Deviation function $\dot{D}Z$ (*cf.* equation 35). Experimental data (16): ♀, 13.61 atm.; O, 40.83 atm.; ȯ, 68.04 atm. Curves: equation 36.

A test of equations 14 and 27 is shown in table 3. The fugacities of hydrogen and nitrogen and the activity coefficients φ_1/φ_1^0 and φ_2/φ_2^0 of their mixtures have been computed from experimental data by Bolshakov (7). His results for 200°C. cannot be quite accurate, since they are not in complete accord with the Duhem–Margules equation. For other gases, especially hydrocarbons, smaller deviations may be expected according to the results reported in table 1.

Since equation 27 is strictly derived from the assumed equation of state and equations 15 and 18, its validity depends only on the validity of these relations. Figure 4 is presented as an example for Z. The maximum deviations for three mixtures are given in table 1.

Experimental P–V–T data for the pure components, if available, can be used for improving the results obtained by means of equation 27. As in equation 13, we represent the difference between the experimental and calculated values of Z_j for the component j by the function W_j. If no data for the mixtures are available the best assumption for the deviation function W of the mixture is

$$W = y_j W_j \tag{28}$$

The improved value of $\ln \varphi_r$ is then obtained by adding

$$\int_0^P W_r \, dP/P \tag{29}$$

to the value obtained by means of equation 27.

If data are available also for mixtures, equation 28 can be replaced by a better interpolation function. The calculation of the additional term in $\ln \varphi_r$ according to equations 25 and 26 will not present any difficulty.

IV. APPROXIMATIONS FOR MODERATE PRESSURE

In the pressure range which is sufficiently well represented by the second virial coefficient (equations 19 and 20), equations 22 and 23 furnish

$$\ln \varphi_r = [B_r - A_r^2 + (A_r - A)^2]P \tag{30}$$

This relation is useful in the calculation of the correction for the imperfection of the vapor in liquid–vapor equilibria. The equilibrium between a liquid (mole fraction x_r, molal volume V'_r and activity coefficient γ'_r of the component r) and its vapor is determined by the condition

$$\ln k_r x_r \gamma_r + V'_r P/RT = \ln y_r \varphi_r P \tag{31}$$

The proportionality factor k_r is eliminated by means of the condition for the equilibrium of the pure component under the vapor pressure p_r^0.

$$\ln k_r + V'_r p_r^0/RT = \ln \varphi_r^0 p_r^0 \tag{32}$$

The equilibrium condition, therefore, is

$$\ln (x_r \gamma_r p_r^0 / y_r P) = (B_r - A_r^2 - V'_r/RT)(P - p_r^0) + (A_r - A)^2 P \tag{33}$$

For binary solutions, the last term becomes $(A_1 - A_2)^2(1 - y_r)^2 P$. A relation of this type has been previously proposed by Scatchard and Raymond (28).

The fugacity rule of Lewis leads to omission of the last term. But this term has in principle the same order of magnitude as the preceding one. However, it is small if the values of the critical constants of the components do not differ appreciably.

In equation 30 the term $(B_r - A_r^2)P$ represents the logarithm of the fugacity coefficient of the pure gas, while $(A_r - A)^2P$ indicates its dependence on the composition. Thus if Z_r for the pure gas is known, we obtain a better approximation by means of the relation

$$\ln \varphi_r = \int_0^P (Z_r - 1)\mathrm{d}P/P + (A_r - A)^2 P \tag{34}$$

in which the equation of state is used only for determining the dependence on the composition.

Another approximation furnishes appreciably better results. We discuss only binary systems, although the method can be extended to the general case without difficulty. We define

$$\dot{D}Z = Z - y_1\dot{Z}_1 - y_2\dot{Z}_2 \tag{35}$$

where $\dot{Z}_1$ and $\dot{Z}_2$ denote the values of Z_1 and Z_2 of the pure components at the pressures y_1P and y_2P, respectively. By means of equations 2 and 6 we can develop the quantity $\dot{D}Z$ in terms of powers of P. The first-order term is

$$\dot{D}Z = y_1y_2P(B_1 + B_2 - 2A_1A_2) \tag{36}$$

This approximation holds quite well up to pressures of about 50 atm. (see figure 5). From equation 36 we derive the corresponding approximation for the fugacity coefficient:

$$\ln \varphi_1 = \int_0^{y_1P} (Z_1 - 1)\,\mathrm{d}P/P + y_2(\dot{Z}_1 - \dot{Z}_2) + y_2^2P(B_1 + B_2 - 2A_1A_2) \tag{37}$$

This quite convenient approximation has an important practical advantage. Sometimes we want to know the fugacity coefficient of a substance in a gaseous mixture below its critical temperature under a total pressure which is higher than its vapor pressure. A relation such as equation 34 cannot be used in this case, because Z_r cannot be measured up to the total pressure P. But equation 37 requires the knowledge of the compressibility factor of the pure gas only up to its partial pressure and therefore is always applicable.

V. CONCLUSION

A variety of problems arise in the calculation of fugacities of gaseous mixtures for two reasons: First, the available experimental data are widely different. Sometimes only the critical temperature and pressure are known, sometimes $P-V-T$ data for some or all of the components, sometimes a complete set of data for the mixtures. In each case, full use should be made of the data. Second, short methods are desirable for moderate pressures. It is believed that the relations discussed in the preceding sections cover the whole range and that the methods proposed do not require an unreasonable amount of time.

Whenever a complete set of data is not available, some assumptions must be introduced which can be only imperfect approximations.

The methods proposed contain only two assumptions: the equation of state (equation 1) and the dependence of the coefficients A and B on the composition (equations 15 and 18).

The equation of state furnishes a second virial coefficient which is in good accord with experimental data at the critical temperature and above. Similarly, the ratio between the limiting volume at high pressure and the critical volume is equal to the average of experimental data which vary within fairly close limits. The equation of state is therefore in accord with general experimental experience to the extent to be expected of an equation containing only two coefficients.

The assumed dependence of the coefficients A and B on the composition of the gaseous mixture takes into account the relative importance of these coefficients at low and high pressures, the theoretical interpretation of these coefficients, and experimental tests of this interpretation.

The degree of approximation obtained by the two assumptions, as ascertained in a number of test examples, is believed to be satisfactory in view of the fact that no experimental data are used except the critical pressure and temperature.

REFERENCES

(1) Amagat, E. H.: Ann. chim. phys. [5] **19,** 384 (1880); Compt. rend. **127,** 88 (1898).

(2) Bartlett, E. P., Cupples, H. L., and Tremearne, T. H.: J. Am. Chem. Soc. **50,** 1275 (1928).

(3) Beattie, J. A., and Bridgeman, O. C.: J. Am. Chem. Soc. **49,** 1665 (1927).

(4) Beattie, J. A., and Ikehara, S.: Proc. Am. Acad. Arts Sci. **64,** 127 (1930).

(5) Beattie, J. A., Kay, W. C., and Kaminsky, J.: J. Am. Chem. Soc. **59,** 1589 (1937).

(6) Benedict, M., Webb, G. B., and Rubin, L. C.: J. Chem. Phys. **8,** 334 (1940); **10,** 747 (1942).

(7) Bolshakov, P.: Acta Physicochim. U.R.S.S. **20,** 259 (1945).

(8) Fowler, R. H., and Guggenheim, E. A.: *Statistical Thermodynamics,* pp. 297 ff. Cambridge University Press, New York and Cambridge (1939).

(9) Gillespie, L. J.: J. Am. Chem. Soc. **48,** 28 (1926).

(9a) Gillespie, L. J.: Phys. Rev. **34,** 352 (1929).

(10) Gilliland, E. R.: Ind. Eng. Chem. **28,** 212 (1936).

(11) *International Critical Tables,* Volume III, p. 9. McGraw-Hill Book Company, Inc., New York (1928).

(12) Joffe, J.: Ind. Eng. Chem. **40,** 1738 (1948).

(13) Kuenen, J. P.: *Die Eigenschaften der Gase* (*Handbuch der allgemeinen Chemie,* Vol. 3), p. 259. Akademische Verlagsgesellschaft, Leipzig (1919).

(14) Lurie, E., and Gillespie, L. J.: J. Am. Chem. Soc. **49,** 1146 (1927).

(15) Michels, A., de Gruyter, J., and Niesen, F.: Physica **3,** 346 (1936).

(16) Michels, A., and Nederbragt, G. W.: Physica **6,** 656 (1939).

(17) Michels, A., and Goudeket, M.: Physica **8,** 347 (1941).

(18) Michels, A., and Geldermans, M.: Physica **9,** 967 (1942).

(19) Morris, W. M., Sage, B. H., and Lacey, W. N.: Am. Inst. Mining Met. Engrs., Tech. Publ. No. **1128** (1939).

(20) Neusser, E.: Physik. Z. **35,** 738 (1934).

(21) Olds, R. H., Sage, B. H., and Lacey, W. N.: Ind. Eng. Chem. **37,** 1008 (1942).

(22) Olds, R. H., Reamer, H. H., Sage, B. H., and Lacey, W. N.: Ind. Eng. Chem. **35,** 922 (1943).

(23) Olds, R. H., Reamer, H. H., Sage, B. H., and Lacey, W. N.: Ind. Eng. Chem. **36**, 282 (1947).
(24) Prigogine, I., and Defay, R.: *Thermodynamique chimique*, Vol. 1, p. 227. Editions Desoer, Liège (1944).
(25) Reamer, H. H., Olds, R. H., Sage, B. H., and Lacey, W. N.: Ind. Eng. Chem. **36**, 88 (1944).
(26) Reamer, H. H., Olds, R. H., Sage, B. H., and Lacey, W. N.: Ind. Eng. Chem. **36**, 956 (1944).
(27) Sage, B. H., and Lacey, W. N.: Ind. Eng. Chem. **26**, 1218 (1934).
(28) Scatchard, G., and Raymond, C. L.: J. Am. Chem. Soc. **60**, 1278 (1938).

LIGHT SCATTERING AS A MEASURE OF PARTICLE SIZE IN AEROSOLS[1]

The Production of Monodisperse Aerosols

DAVID SINCLAIR[2] AND VICTOR K. LA MER

Department of Chemistry, Columbia University, New York, New York

Received October 18, 1948

INTRODUCTION

The measurement of colloid particle size by observation of the scattered light is based primarily on the electromagnetic theory developed by Gustav Mie (16) in 1908. Previously, Rayleigh (18) had derived his well-known equations for scattering by small particles of radius much less than the wave length of the incident light. Since that time, additional investigations, both theoretical and experimental, have been made by Debye (3), Love (15), Blumer (1), Jobst (7), Engelhard and Friess (4), Ray (17), Bromwich (2), Stratton and Houghton (22), Ruedy (19), La Mer and Barnes (10), Johnson and La Mer (8), La Mer and Kenyon (12; see also 9, 11), Van de Hulst (23), and others. The complete theory is given in a compact form by Stratton (21). The pertinent equations of the Mie theory follow.

Consider a plane wave of light incident on any isotropic dielectric sphere of radius, r, small compared to the wave length. The total light energy scattered per second per unit intensity of illumination (unit energy per unit cross-sectional area) is:

$$S = 24\pi^3 \left|\frac{m^2 - 1}{m^2 + 2}\right|^2 \frac{V^2}{\lambda^4} \tag{1}$$

Here V is the volume of the sphere equal to $4\pi r^3/3$, λ is the wave length of the light in the surrounding medium, and m is the real or complex refractive index of the particle relative to the real refractive index of the medium (see appendix A). S is defined as the scattering cross-section of the sphere. Equation 1 is in the form given by Rayleigh (18) and it applies only when $r << \lambda$.

For spheres of any size the general form of equation 1 is:

$$S = \frac{\lambda^2}{2\pi} \sum_{n=1}^{\infty} (2n + 1)\left(|a_n|^2 + |p_n|^2\right) \tag{2}$$

---

[1] Presented at the Symposium on Aerosols which was held under the joint auspices of the Division of Physical and Inorganic Chemistry and the Division of Colloid Chemistry at the 113th National Meeting of the American Chemical Society, Chicago, Illinois, April 22, 1948.

This article is based in part on work performed for the Office of Scientific Research and Development under Contract NDRC-r33 and OEMsr-148 with Columbia University, in 1940-42. Recommendation for the contract was made by the Aerosol Section of NDRC (Dr. W. H. Rodebush, Chief) and the work was carried out under the supervision of this Section.

[2] Present address: Johns-Manville Research Center, Manville, New Jersey.

The functions a_n and p_n are complex functions of the parameter $\alpha = 2\pi r/\lambda$ and the index of refraction m. For transparent (non-absorbing) spheres m is real and equation 2 yields the total amount of light abstracted (i.e., scattered) from the incident beam. For absorbing spheres m is complex and equation 2 yields only that part of the light which is scattered by the sphere. (See appendices A and B for the meaning of complex index of refraction and the definitions of a_n and p_n.)

For absorbing spheres the total light abstracted from the incident beam (i.e., both scattered and absorbed) per second per unit intensity of illumination is:

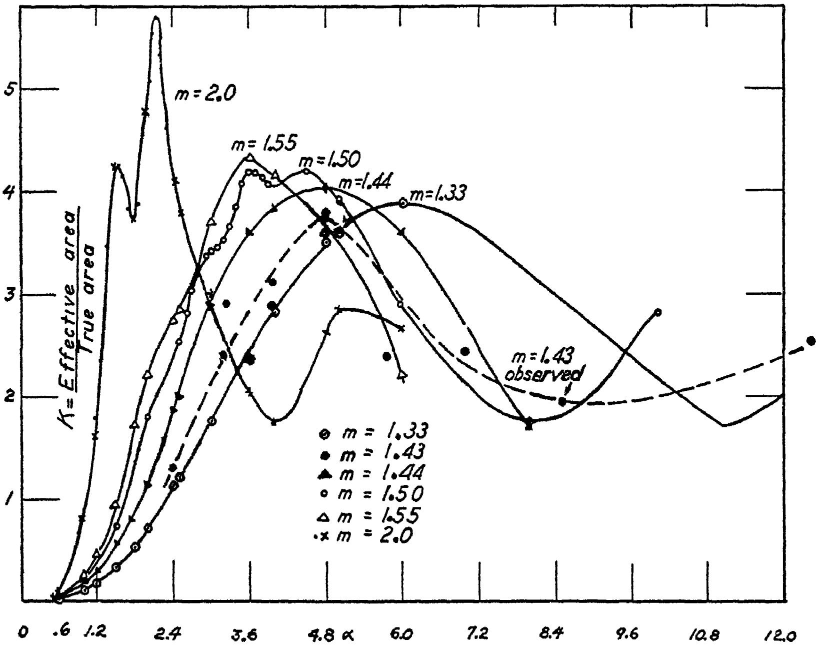


FIG. 1. Total scattering, K_s, per unit cross-sectional area of particle *versus* α (calculated). $m = 1\ 33, 1.44, 1.50, 1.55, 2\ 00$; observed $m = 1.43$.

$$S_e = \frac{\lambda^2}{2\pi} Re \left[\sum_{n=1}^{\infty} (2n + 1)\ (a_n + p_n) \right] \tag{3}$$

S_e is defined as the extinction cross-section and Re stands for the real part of the expression in the brackets.

The sum in equation 2 has been calculated for a variety of values of α and m (real) by The Mathematical Tables Project of the Bureau of Standards. Tables of these calculations have recently been published (14). Figure 1 shows the results obtained for five values of the real refractive index m. The ordinates are the "scattering area coefficient", defined as $K_s = S/\pi r^2$, i.e., the ratio of the scattering cross-section to the geometric cross-section. The abscissas are the values of $\alpha = 2\pi r/\lambda$. From equation 2 it is seen that

$$K_s = \frac{2}{\alpha^2} \sum_{n=1}^{\infty} (2n + 1) \left(|a_n|^2 + |p_n|^2\right) \tag{4}$$

so that for a given refractive index K_s is a function of α alone.

The Mathematical Tables Project has also calculated the extinction cross-section, S_e, from equation 3, for a variety of absorbing particles. These calculations have not as yet been published. (See reference 8a on the scattering of ultraviolet light by sulfur hydrosols.)

The above equations hold for incident light of any state of polarization. However, the intensity scattered in a given direction varies with the state of polarization of the incident light. For example, in the limiting (Rayleigh) case when the particles are very small compared to the wave length and the *incident light is plane polarized,* the intensity scattered at an angle ψ when $r < < \lambda$ is:

$$I_\psi = \frac{9\pi^2}{R^2} \left|\frac{m^2 - 1}{m^2 + 2}\right|^2 \frac{V^2}{\lambda^4} \sin^2 \psi \tag{5}$$

R is the distance from the sphere to the point of observation, which must be very large compared to the radius of the sphere. ψ is the angle between the direction of propagation of the scattered light and the electric vector in the incident light. The scattered light is plane polarized, with its electric vector lying in the plane defined by the electric vector in the incident light and the direction of propagation of the scattered light.

When the incident light is unpolarized, and $r < < \lambda$, the intensity scattered in the direction θ is:

$$I_\theta = \frac{9\pi^2}{2R^2} \left|\frac{m^2 - 1}{m^2 + 2}\right|^2 \frac{V^2}{\lambda^4} (1 + \cos^2 \theta) \tag{6}$$

Here θ is the angle between the directions of propagation of the incident and scattered light.

The theory predicts and observation confirms that when the incident light is unpolarized the light scattered by spheres of any size is composed of *two incoherent, plane polarized components* whose planes of polarization are mutually perpendicular. For the extremely small spheres to which equation 6 applies, the $\cos^2$ term in the parentheses gives the relative intensity of the scattered component whose electric vector lies in the plane defined by the incident beam and the observed scattered beam (hereafter referred to as the plane of observation). The term corresponding to the factor unity in the parentheses refers to the scattered component whose electric vector is perpendicular to the plane of observation. When $\theta = 90°$ it is seen that the scattered light is plane polarized with its electric vector perpendicular to the incident beam, a well-known result.

Equation 5 (but not equation 6) is in the form given by Rayleigh (18), who omitted the factor $\frac{1}{2}$ from equation 6. It should be emphasized that both equations refer to unit intensity of illumination. Our experimental and theoretical investigations have shown that the value usually given in the literature for the

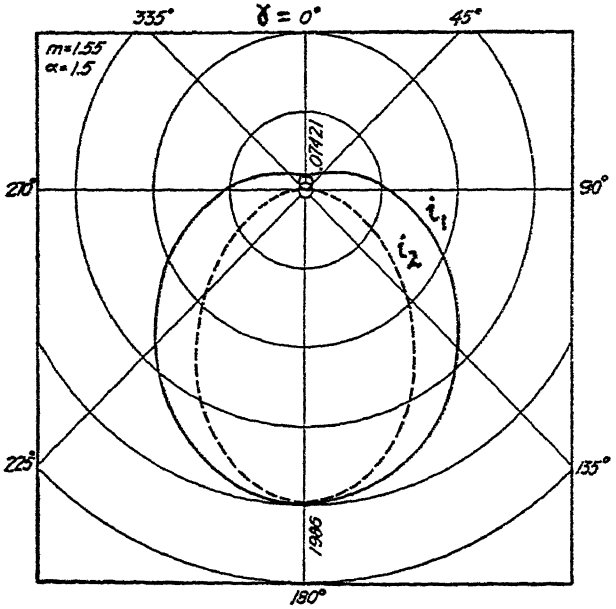


FIG. 2. Polar diagram of i_1 and i_2 for $m = 1.55$ and $\alpha = 1.5$

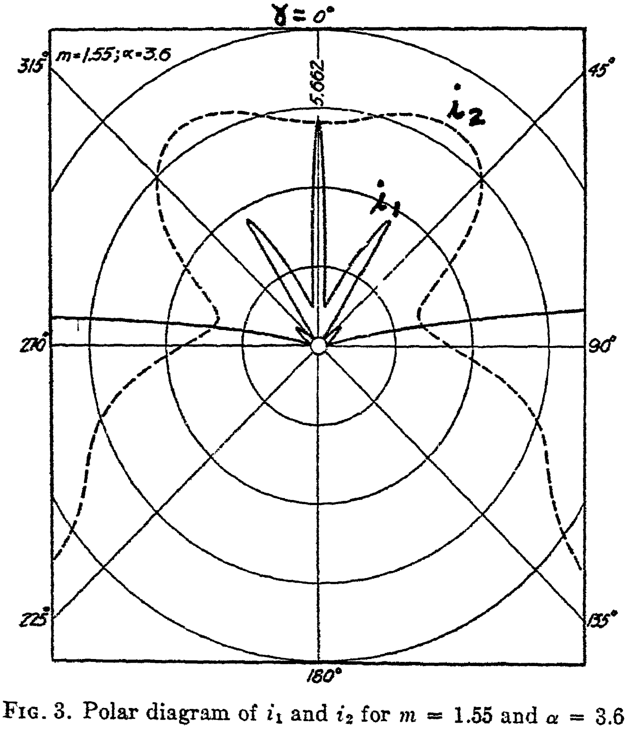


FIG. 3. Polar diagram of i_1 and i_2 for $m = 1.55$ and $\alpha = 3.6$

intensity scattered by particles illuminated by unpolarized light is too large by a factor of two, for further reasons explained elsewhere (14, 20).

For spheres of any size the general form of equation 6 should be given by Mie as:

$$I_\gamma = \frac{\lambda^2}{8\pi^2 R^2}(i_1 + i_2) \tag{7}$$

(Mie writes the factor as $\frac{1}{4}$ but it should be $\frac{1}{8}$ for the reasons given in references 14 and 20.) The angular distribution functions i_1 and i_2 are complicated functions of α, m, and θ. (See appendix C for definitions of i_1 and i_2.) Their values have been calculated by The Mathematical Tables Project for several values of α and m (real), for values of γ ($\gamma = 180° - \theta$) from 0° to 180° in steps of 10° (14). Typical results are shown in figures 2 and 3, where polar diagrams of the functions i_1 and i_2 are plotted for $m = 1.55$ and $\alpha = 1.5$ and 3.6. Similar diagrams for smaller values of m are plotted in reference 1.

The functions i_1 and i_2 are proportional to the intensities of the two incoherent, plane polarized components in the scattered light. i_1 is the component whose electric vector is perpendicular to the plane of observation, and i_2 is polarized perpendicular to i_1. When the sphere is illuminated by unpolarized light of unit energy per unit cross-sectional area, the scattered intensity is given by equation 7. (When the incident light is plane polarized, the scattered light is elliptically polarized and i_1 and i_2 do not in general give its components.)

Theoretically the above equations can be used to measure the radius of spherical particles of any size composed of either transparent or absorbing materials. In the work described below the measurements were limited to aerosols of spherical transparent droplets, since calculations for absorbing materials were not then available. The work is further limited to radii below about 1 micron because of the practical difficulties of evaluating equations 2 and 7 for larger sizes. (Measurements of absorption in liquid sulfur sols will be found in references 8, 9, 11, and 12.) Furthermore, above 1 micron the ordinary light microscope becomes increasingly accurate and useful and accordingly the more precise methods based upon scattering become less important.

The measurements were of four types: (*1*) observation of the transmitted light as a function of wave length; (*2*) the intensity of the scattered light at a given wave length; (*3*) the color of the scattered light; (*4*) the polarization of the scattered light at a given wave length. Precise measurements were made on aerosols of uniform droplet size produced by a method described in a later section.

Measurements were also made on fogs containing a distribution of sizes. Such measurements yield an average value of the radius whose accuracy depends on the degree of uniformity of size. In many practical cases, such as the oil fog from an outdoor screening smoke generator or the aerosol required for filter penetration tests, the distribution of sizes is comparatively narrow so that the equations applicable to perfectly uniform fogs can be used with fair accuracy.

Measurements of the size distribution in non-uniform fogs were made by combining gravity settling methods with the optical methods.

The concentration was low enough so that the scattering was *incoherent* and the effect of each of the n droplets per unit volume was superposed. This occurs when the ratio of droplet radius to distance of separation is about 100, corresponding to a fog of 10^6 droplets per cubic centimeter and of radius 1 micron. The mass concentration of such a fog is about 4 mg. per liter. The mass concentration usually occurring in the laboratory and in oil fogs from an outdoor generator is a few tenths of a milligram per liter or less.

1. METHODS BASED ON TRANSMISSION

Transmission measurements

For transparent materials the scattering cross-section, S (equation 2), can be obtained by measuring the transmission as a function of wave length. Measurements at various concentrations showed that the transmission, T, could be expressed by the equation:

$$T = \frac{I}{I_0} = e^{-Snl} = e^{-Jcl} \tag{8}$$

Here I_0 is the incident and I the emergent intensity of light of wave length λ, n the particle number concentration, c the mass concentration, J the scattering cross-section per gram, and l the length of path in the aerosol. When either the concentration or the radius is known, the other can be found by measuring the transmission at a known wave length. When both are unknown, they may be found by measuring the transmission at two or more wave lengths.

A beam of approximately monochromatic light from a Hilger Monochromator was passed through a uniform particle size aerosol of known concentration. The transmission was measured with a photocell, as a function of particle size and wave length, and the scattering cross-section per gram calculated from equation 8. Measurements were made at values of T from about 0.9 to 0.3 for a path length of 1 m. Good agreement was obtained between the experimental values of J and the theoretical values for oleic acid ($m = 1.46$), petroleum-diol ($m = 1.55$), and sulfur ($m = 2.0$). The results for oleic acid are given in figure 4. The ordinates λJ are a function of α alone, and therefore valid for all radii and wave lengths, and the abscissas are log r/λ (theoretical) and $1/\lambda$ (experimental). The curves *a* and *b* show the experimental results obtained for oleic acid of refractive index $m = 1.46$ and the curve to the left is the theoretical curve for $m = 1.44$. The maxima for the experimental curves occur at $\lambda J = 5.0$ cm.^{3} per gram of oleic acid, in satisfactory agreement with the theoretical maximum of 5.4 cm.^{3} per gram. The droplet radius determined from the horizontal displacement of the experimental maximum from the theoretical is 0.38 micron and 0.29 micron for curves *a* and *b*, respectively, in satisfactory agreement with the

---

[3] The transmission method, applied to measurements on liquid sulfur hydrosols, is described in greater detail in references 8, 9, 10, 11, and 12.

observed higher-order Tyndall spectra ($3\frac{1}{2}$ and $2\frac{1}{2}$ orders, respectively) described below. The measurements on oleic acid and sulfur are described in reference 14.[3]

This method was also applied in field measurements by means of a portable photoelectric meter developed by Dr. Seymore Hochberg, called the "slope-o-meter." The relative transmission at two wave lengths through a sample of smoke gave an average particle radius, and the absolute transmission at one wave length gave the concentration.

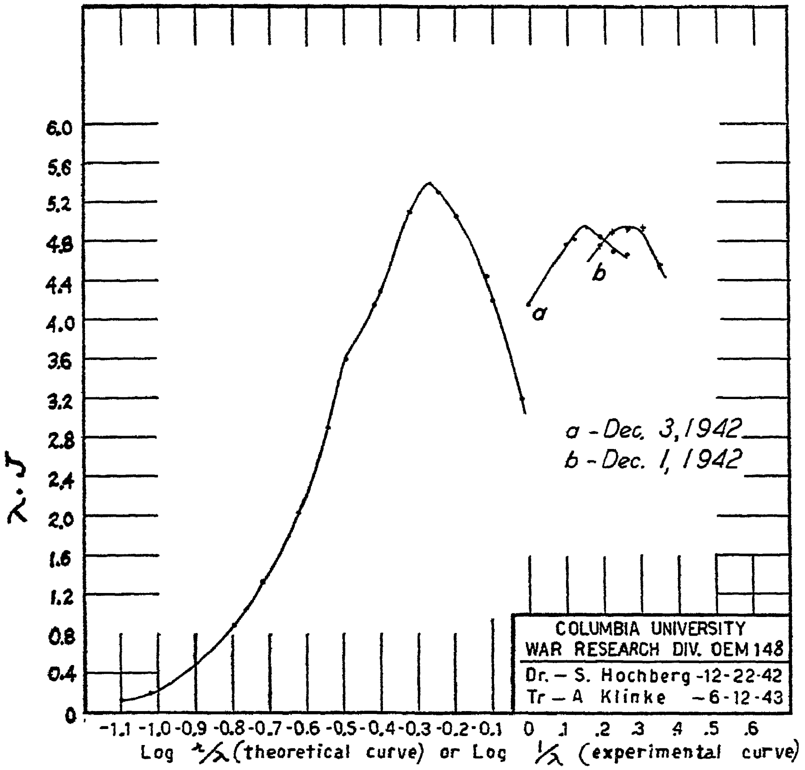


FIG. 4. Comparison of experiment with Mie theory. Oleic acid smoke. The product of wave length λ (cm.) and scattering coefficient J (cm.^{2} per gram) as a function of radius and wave length. The experimental curves a and b exhibit a close approach to the theoretical value of $\lambda J = 5.4$ at the maximum.

Visual observations

Visual observation of the residual color of white light transmitted through a fog provides a convenient and rapid measure of the average droplet size. Referring to the scattering coefficient K in figure 1, where the abscissa α is proportional to r/λ, i.e., inversely proportional to wave length, it will be seen that for the values of α to the left of the maximum, the *scattering decreases as the wave length increases.* In other words, the transparency increases with increasing wave length. If the sun is observed through a fog whose droplet radii are smaller than the radius corresponding to the principal maximum, the sun's disc will appear a deep red color just before extinction. *Conversely, if the radii are greater than the maximum in the K curve of figure* 1, *the residual color will be blue or green.* If the radii are grouped closely about the maximum, the residual color will be magenta. When the spread in particle size is too great, i.e., about as much as the wave length ratio of red to violet light, little or no color will be seen. This method was used to estimate the average droplet size and size distribution in the oil fogs produced outdoors. The results were in satisfactory agreement with those obtained from the photoelectric transmission meter and a portable meter (called the Owl) for scattered light described below.

2. METHODS BASED ON THE INTENSITY OF THE SCATTERED LIGHT

Absolute intensity

The radius of the particles of a fog of uniform droplet size may be obtained by measuring the intensity scattered at angles θ varying from the forward to the backward direction and integrating over a sphere. The method is tedious and the results are not as accurate as transmission measurements.

A fog of supercooled stearic acid droplets of known uniform radius and concentration was observed while streaming out of a generator, in the form of a cylindrical jet about 1 cm. in diameter. A suitable volume was uniformly illuminated by unpolarized white light and the intensity scattered into a given small solid angle was measured with a photometer. The measurements were made through a roughly monochromatic filter ($\lambda = 0.524 \pm 0.01$ micron) at angles θ from 3° to 175°, at frequent intervals. The intensities were then integrated over a sphere and the scattering cross-section, S, obtained by comparison with the brightness of a diffuse reflector of known reflectivity. The results are given in the dashed curve of figure 1, in satisfactory agreement with the theoretical curve for $m = 1.44$.

The integration over the sphere of the observed scattered light when the incident light is unpolarized is mathematically equivalent to the integration of equation 7, which should be equal to equation 2. It was found, however, that the integral of equation 7 with the factor $\frac{1}{4}$ as given by Mie was equal to twice the value of S given by equation 2. The satisfactory agreement between the experimental values and equation 2 shows that the absolute intensities i_1 and i_2 as originally given by Mie for equation 7 are twice too high, as mentioned above. The detailed explanation has been given elsewhere (20).

Relative intensity

(a) Fogs of non-uniform droplet size

At a given wave length the observed intensity scattered at any angle, θ, by particles of a given radius is directly proportional to the concentration. This is true provided the concentration is low enough so that secondary scattering is negligible. In this case the relative intensities scattered by particles of two different radii may be obtained from equation 7. For this purpose the intensities calculated from equation 7 were approximated over a small range of values of the radius as proportional to some power of the radius. Thus the intensity, I, scattered in a given direction from a given volume of aerosol is proportional to the particle number concentration, n, and the particle radius raised to some power, i.e.,

$$I = kr^p n \tag{9}$$

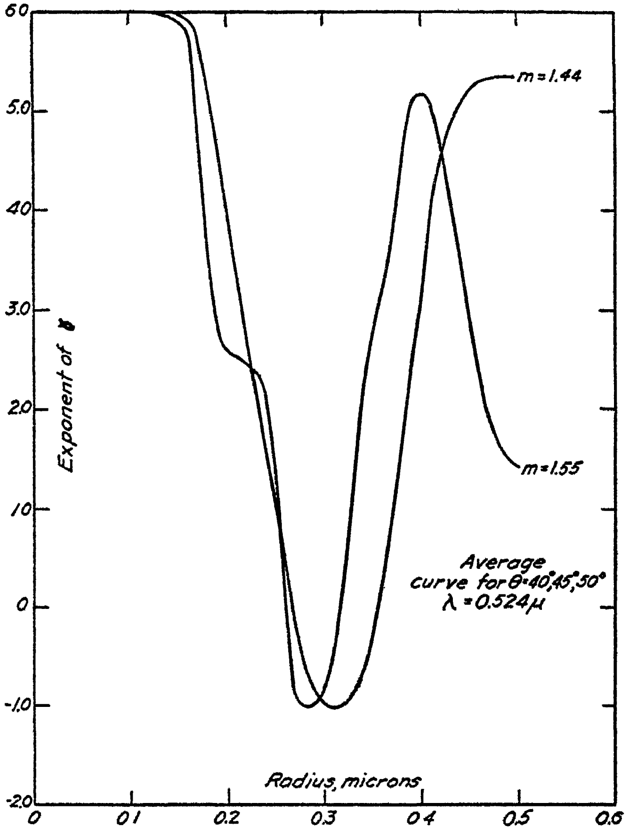


FIG 5. Variation of p with r at $\theta = 40°, 45°, 50°$

The value of p as a function of r was obtained from the values of $(i_1 + i_2)$ given in the tables (14). The values of $\log (i_1 + i_2)$ *vs.* $\log r$ were plotted and the values of the slope, p, as a function of r obtained for several angles.

Figure 5 shows the variation of p with r at angles $\theta = 40°, 45°$, and $50°$ (averaged), for refractive indices 1.44 and 1.55 at the wave length $\lambda = 0.524$ micron. It is seen that p has the value 6 for radii below 0.1 micron, as is to be expected in the Rayleigh region. There is also a *region where the slope is negative,* that is, the scattering per particle actually decreases with increasing particle radius. This region, of course, corresponds to the one mentioned in the section on visual observations, where for a constant radius the *scattering increases* with increasing wave length, so that a blue transmission prevails.

The data in figure 5 may be used to obtain the size distribution in a fog of non-uniform droplet size by allowing it to settle in a convection-free fog chamber and observing the decrease in scattered light intensity with time. The apparatus consisted of a cylindrical brass chamber 2½ in. high and 3 in. in diameter submerged in a thermostated water bath. Fog was introduced through tubes extending down through the water bath into the fog chamber, and illuminated and observed through plane glass windows. As the fog settled, the intensity scattered by a small volume at a fixed height was measured as a function of time. The concentration of the fog was low enough so that multiple scattering and coagulation were negligible.

Consider a layer in the fog at a height x below the top of the chamber, and suppose that at time $t = 0$ the concentration is uniform throughout the chamber. At the time $t_1 = x/v_1$ there will be no droplets above this layer whose velocity of fall is greater than v_1, and at a later time $t_2 = x/v_2$ there will be no droplets above this layer whose velocity of fall is greater than v_2, where v_2 is less than v_1. Consequently, observation of the decrease in number concentration at a height x during the time interval $t_2 - t_1$ will give the number of particles having velocities of fall between v_2 and v_1, or radii between r_2 and r_1 given by Stokes' law of fall.

The decrease in number concentration is given by equation 9 in terms of the decrease in scattered light intensity, that is,

$$\Delta n = \Delta I/kr^p \tag{10}$$

where r is now the average of the radii r_1 and r_2. The corresponding value of p is obtained from figure 5. Successive measurements at different times will give the relative number of droplets of different radii.

Observations were made on an illuminated volume of fog at an angle of 45° to the incident beam. The observed volume was located a distance $x = 1.5$ cm. below the top of the chamber and the height of this volume was 2 mm. With these dimensions the intensity observed at any one time is that due to all droplets lying within the layer 1.5 ± 0.1 cm. from the top. This introduces an error such that a fog of uniform droplet size appears to have a spread in size of 7 per cent.

In case the law of scattering as a function of droplet size is not known for a given apparatus, it may be calibrated by observation of the intensity scattered by uniform droplet size fogs of known concentration, as described in the next section.

(b) Fogs of uniform droplet size

When the fog is made up of droplets of uniform radius the size may be obtained by observing the rate of settling of the top of the cloud in a convection-free settling chamber. The apparatus used was similar to that described above except that the smoke chamber was a glass cylinder about 3 in. in diameter and 1 ft. high immersed in a larger glass cylinder through which water from a thermostat was circulated. The height of the cloud was recorded at suitable time intervals by sighting along the top of the cloud and recording the position on the wall of the glass water jacket.

The concentration corresponding to an observed scattered intensity was obtained by passing a known volume of aerosol through a glass wool filter and weighing. The Mie theory shows that the intensity scattered at angles $\theta = 5°$ to 30° may be 100 to 1000 times that scattered at 90°. Hence, the intensity measurements were made with a new form of Tyndall meter which collected and brought to a focus all the light scattered at small angles from the forward direction. The aerosol to be collected in the glass wool filter was first passed through a cylindrical chamber 3 in. in diameter and 4½ in. long. A small volume of the aerosol at the center of the chamber was intensely illuminated by means of a 16-cp. automobile headlight bulb. The image of the filament was formed at this point by two aspheric condenser lenses, 2½ in. in diameter and of 2¼ in. focal length, located at one end of the chamber. The central portion of the wide angle cone of rays was blocked out by a disc ½ in. in diameter, and the intensity of the smoke was observed through the other end of the chamber along the axis of the cone. The diameter of the observation window was slightly less than that of the blocked-out region so that no direct rays reached the photometer. An opaque screen in the center of the chamber, having an aperture just larger than the image, cut off the light scattered from the lens surfaces.

When no particles are present, the observer sees only the black background of the disc. When aerosol is passed through, the background becomes brilliantly illuminated by the light scattered at 5–30° from the direction of the incident cone of rays. The photometer is focussed on the image of the filament formed in the smoke. The brightness of this image can be measured quite accurately except when the particle size is so large and the concentration so low that flickering occurs. The sensitivity of the apparatus is such that the dust particles in the ordinary air of a room appear brilliantly illuminated. Consequently, the chamber must be air-tight and the zero reading made with well-filtered air.

This method has been further refined by Hochberg in this laboratory, and later by Gucker (5) and collaborators at Northwestern University. They measured the scattered light by using photoelectric cells in place of the visual photometer. The calibration requires the aerosols of uniform particle size which are described in the next section and are mentioned in Gucker's publication. The instrument was calibrated for smoke of a given particle size by observing the intensity at several concentrations high enough to be measured accurately by collection in the glass wool filter, but low enough so that secondary scattering and coagulation were negligible. The concentration of more dilute smokes is then directly proportional to the intensity.

This method was used for measurements of mass concentration from 200 down to 10^{-3} micrograms (10^{-9}g.) per liter, depending upon the particle size. It is believed to be the most sensitive aerosol detector ever reported. It was equalled, but not exceeded, in delicacy by the later more complicated instruments developed by Gucker and associates. For all particle sizes, the secondary scattering was negligible for concentrations of 200 micrograms per liter or less. However, coagulation sets an upper limit to the allowable mass concentration, which decreases with particle size. For example, a smoke of 0.5 micron radius will not show appreciable coagulation at concentrations up to 500 micrograms per liter, while the concentration of a smoke of 0.15 micron radius should not be greater than 14 micrograms per liter.

3. METHOD BASED ON THE COLOR OF THE SCATTERED LIGHT: HIGHER-ORDER TYNDALL SPECTRA

Visual observation of the color of the light scattered at different angles by a fog of uniform droplet size provides a rapid measure of particle size. When the fog is illuminated with a parallel beam of unpolarized white light, a series of bright colors is seen as the angle of observation, θ, is varied. The observation is best made through a plane polarizer oriented with its vibration direction perpendicular to the plane of observation, in which case the component i_1 alone is seen. The component i_2 exhibits a different and usually less distinct series of colors.

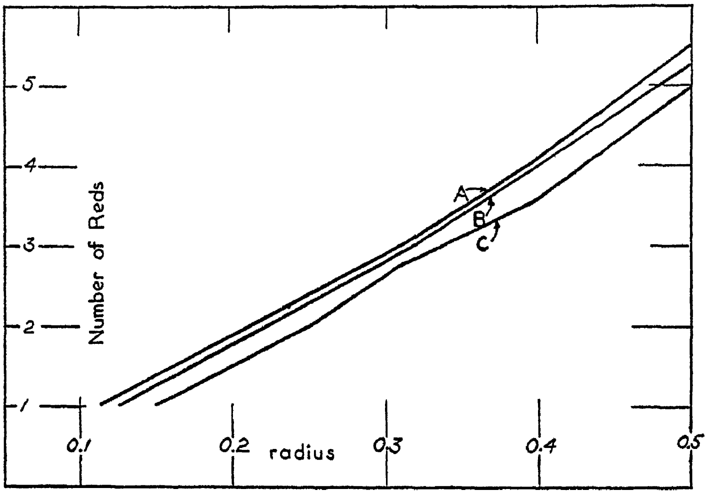


FIG. 6. Number of reds observed plotted against radius. A, sulfur (index = 2.0); B, stearic acid (index = 1.43); C, calculated (all indexes).

As the angle of observation is varied from near 0° toward 180°, the sequence of colors resembles the spectrum in the order violet, blue, green, yellow, orange, red. This spectral sequence is repeated several times depending on the particle size. Near 90° the sequence of the colors is reversed. The spectral sequences have been designated as higher-order Tyndall spectra (8), for which the Rayleigh type would be of order zero. The purity and brightness of the colors increase with uniformity of droplet size. The number of times the spectral series is repeated increases with droplet radius. The droplet radius may be measured with considerable accuracy by counting the number of times red, the most distinctive color, is seen.

Figure 6 shows the experimental curves obtained with stearic acid fogs (m = 1.43) and sulfur fogs (m = 2.00). The ordinates are the number of reds seen when the angle of observation is varied from near 0° to near 180°, and the abscissas are the radii in microns, obtained from measurements of the rate of settling in the convection-free chamber using Stokes' law. It is seen that the radius in microns is roughly equal to the number of reds divided by 10. The number of reds to be expected theoretically was calculated from the values of the angular distribution function i_1 (equation 7) given in the tables (14). A curve was plotted for each particle radius showing the ratio of i_1 at $\lambda = 0.629$ micron (red) to i_1 at $\lambda = 0.524$ micron (green) *versus* scattering angle from 0° to 180°. The number of maxima of ratio greater than 0.45 was taken equal to the number of reds observed visually. The ratio 0.45 was chosen, since the intensity ratio at these two wave lengths is 0.9 in sunlight and 2.0 in the tungsten light used for observation. Consequently, a calculated ratio greater than 0.45 would correspond to an observed reddish hue, since the ratio in the observed scattered light would be greater than $2 \times 0.45 = 0.9$. The tables (14) were designed to permit this calculation of the colors in scattered light by choosing successive pairs of α-values which bear the ratio 1.2, the ratio of 0.629 to 0.524. The results are shown in curve C of figure 6. The number of maxima was found to be practically independent of refractive index throughout the range calculated from the tables. It is seen that the calculated curve differs from the experimental curves by about 0.025 micron.

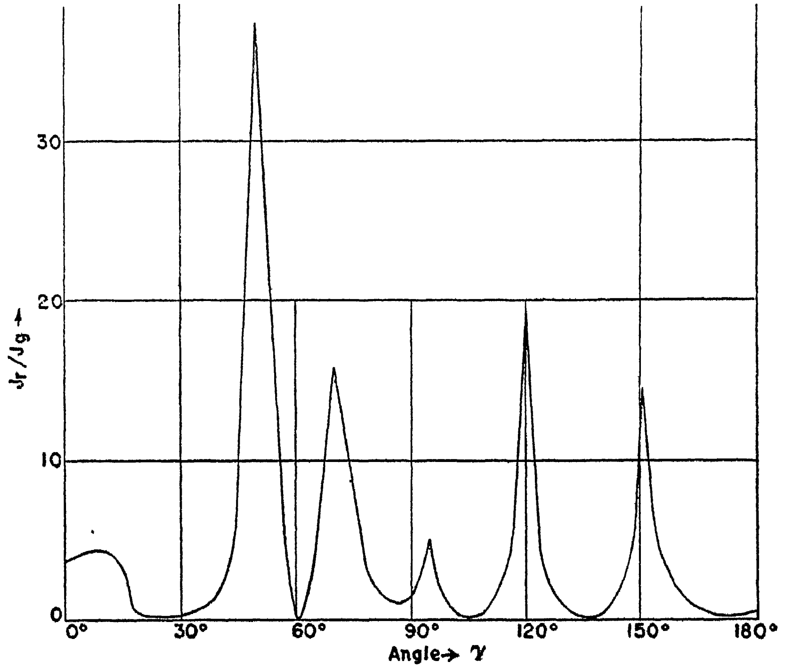


FIG. 7 Number of maxima calculated: i_1 (red) to i_1 (green) *vs.* angle γ. Red-green intensity ratio. $\alpha_r = 5$; $\lambda_r/\lambda_g = 1.2$; $m = 1.44$.

Figure 7 shows the calculated red-green intensity ratio as a function of angle $\gamma = 180° - \theta$ and $m = 1.44$, corresponding to a six-order stearic acid smoke.

Figure 8 shows the angular position of the reds observed in stearic acid fogs. With the exception of sulfur the calculated points for all indices of refraction lie on the stearic acid curves to within 5° of angle. With the aid of figure 8 greater accuracy in the particle-size measurement can be obtained by observing the angular position of the reds[4] as well as their number. The colors may be observed conveniently in a spherical flask of fog held in the hand in the parallel light beam. When the droplet size is very uniform and not larger than about 0.6 micron radius the red orders may be counted without the aid of a plane polarizer. When the droplet size is not very uniform or the radius is greater than 0.6 micron, the colors produced by the different sizes and by the two components i_1 and i_2 overlap and cause confusion. The introduction of a plane polarizer such as a disc of Polaroid between the eye and the flask gives increased clarity. No colors can be seen in fogs having a wide distribution of sizes.

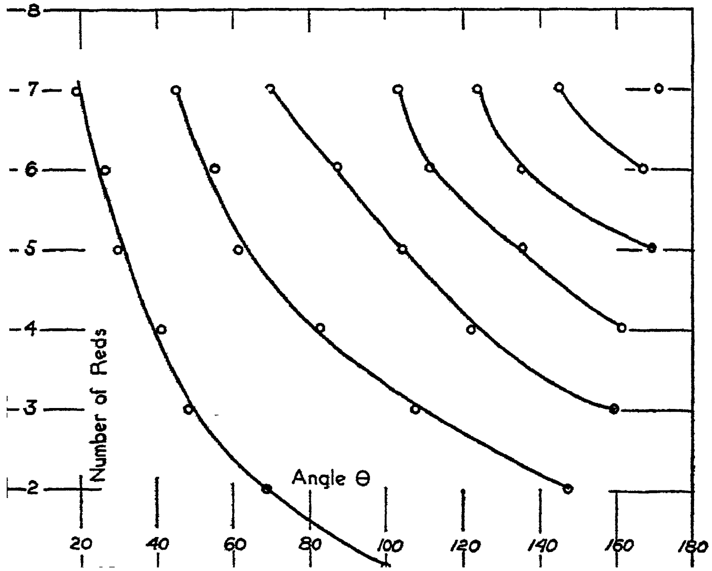


FIG. 8. Stearic acid fogs: angular position of reds observed in scattered light

When the Polaroid is held in the path of the incident light beam the observed scattered light is elliptically polarized except when the vibration direction of the Polaroid is either perpendicular or parallel to the plane of observation. With the first orientation the component i_1 alone and with the second orientation the component i_2 alone appears in the scattered light.

For accurate measurement of the angular position of the colors, the flask of fog was observed through a telescope arranged to rotate about an axis passing through the center of the flask. The light beam and telescope were directed at the center of the flask and perpendicular to the axis. The telescope carried a

---

[4] This method is *analogous* to the measurement of the angular position of the *coronas* observed in the light *diffracted by still larger droplets*. Humphreys (6) gives a calculated curve showing the angular radii of the first and second red coronal rings as a function of the radii of water droplets between 1 and 10 microns. It will be seen that these two curves are roughly extensions of the first two curves of figure 8.

Polaroid which could be oriented perpendicular or parallel to the plane of observation, so that the series of colors produced by either of the components i_1 and i_2 could be observed. It was found that i_1 gave a more distinct series of colors over a greater range of sizes.

A portable instrument (13), suitable for use either in the laboratory or in the field, was constructed for rapid measurement of the number and position of the reds. The instrument (code name "Owl") consisted of a cylindrical smoke chamber and a light source which could be held in the hand and rotated while observing the smoke through a low-power microscope. The axis of rotation was perpendicular to the plane of observation. Smoke could be drawn in and out by means of a rubber bulb and ball valves. A cylindrical window in the side of the chamber allowed observation at angles from 15° to 165° read from a graduated scale. It was found that all the reds could usually be observed in this range. The parallel beam from the light source[5] traversed the cell along a diameter and entered a tube which served as a light trap to reduce stray light. The inside of the chamber was blackened by painting and covering with soot.

The eyepiece of the microscope contained a split-field Polaroid disc of which one half had its vibration direction perpendicular and the other half parallel to the dividing line. The scattered light was observed with the dividing line perpendicular or parallel to the plane of observation, so that the component i_1 was seen in one half and i_2 in the other. The two series of colors produced by i_1 and i_2 were thus observed simultaneously but separately as the chamber and light source were rotated. This arrangement sometimes aided in estimating the size of smokes of non-uniform particle size. By placing a plane polarizer (hereafter referred to as the analyzer) between the eye and the split-field disc, the instrument was converted into a polarization photometer. In this form it was used to measure the relative intensities of the components i_1 and i_2 as described in the next section.

4. METHOD BASED ON THE POLARIZATION OF THE SCATTERED LIGHT

In general the relative intensity of the plane polarized components i_1 and i_2 varies in a very irregular manner with α. However, for values of α below 2.5 and for angles of observation near 90°, the intensity ratio i_2/i_1 is a singly valued function of α and may be used as a measure of particle size.

The polarization photometer referred to above was used to measure the intensity ratio i_2/i_1 at $\lambda = 0.524$ for radii from 0.05 to 0.2 micron. A fog of uniform droplet size was observed through a Wratten green filter placed between the eye and the analyzer. With the dividing line of the split-field disc set perpendicular or parallel to the plane of observation, the analyzer was turned until the intensities of the two halves of the field were equal. If the angle between the light vibrations transmitted by the analyzer and the plane of observation is ϕ, then:

$$i_2/i_1 = \tan^2\phi \tag{11}$$

The values of i_1 and i_2 given in the tables (14) were used to obtain theoretical calibration curves. Figure 9 shows the calculated values of ϕ as a function of radius for five different refractive indices when the angle of observation $\theta = 90°$ and $\lambda = 0.524$ micron. Figure 10 shows ϕ as a function of radius for four values of ϕ when the index of refraction $m = 1.44$ and $\lambda = 0.524$ micron.

---

[5] Such as the Bausch & Lomb Nicholas Illuminator.

The calibration curves of figure 9 were used to measure the droplet radius of fogs of uniform size. Since the size was so small it was not possible to check the calibration curves by other experimental methods except at isolated points. Good checks were obtained for supercooled sulfur droplets of 0.11 micron radius and oleic acid droplets of 0.17 micron radius by observation of the rate of fall of uniform-sized fogs in a small, convection-free settling chamber.

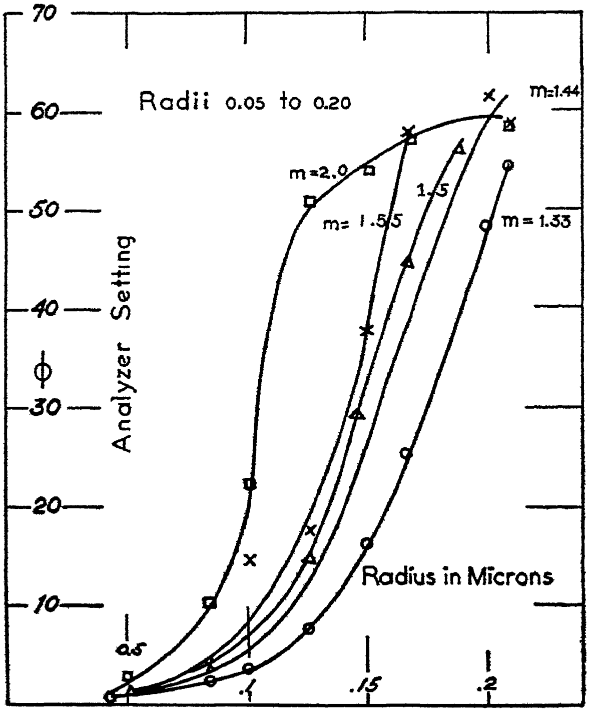


FIG. 9. Calibration curves for the Owl. Analyzer angle ϕ *vs.* r for $\theta = 90°$ at five refractive index values.

Since good agreement with theory was obtained for larger sizes when using the other experimental methods described above, the polarization method is considered to be reliable for fogs of small droplet radii provided they are of sufficiently uniform size. The uniformity of size can be checked by measuring the polarization at two or more of the angles of observation given in figure 10. If the droplet radius obtained from each of these measurements is the same, then the fog is of reasonably uniform size.

It should be emphasized that the polarization method can be used only for small droplet fogs of radius below about 0.2 micron, depending upon the refractive index. In short *the method should not be used when the number of reds observed is two (2) or greater*, as the function is then not monotonic.

METHOD OF PRODUCTION OF FOGS OF UNIFORM DROPLET SIZE (13)

Stable aerosols of very uniform particle size may be produced by slow and uniform condensation of the vapors of liquids of boiling point between 300° and 500°C. when well mixed with air containing condensation nuclei. The size of the particles is determined by the ratio of the mass of condensible vapor to the number of nuclei. When the cooling and other pertinent factors are carefully controlled, it is possible to produce aerosols having a particle radius which does not vary by more than 10 per cent from the average, as shown by the direct microscopic measurement of the droplets and the disappearance of the higher-order Tyndall spectra when aerosols of different particle radii are mixed. The reliability of this criterion of uniformity of particle size has since been tested by mixing sulfur hydrosols of uniform particle size. When sols for which the radii of the particles differed by 2 per cent were mixed, the fifth and sixth orders were obliterated (8).

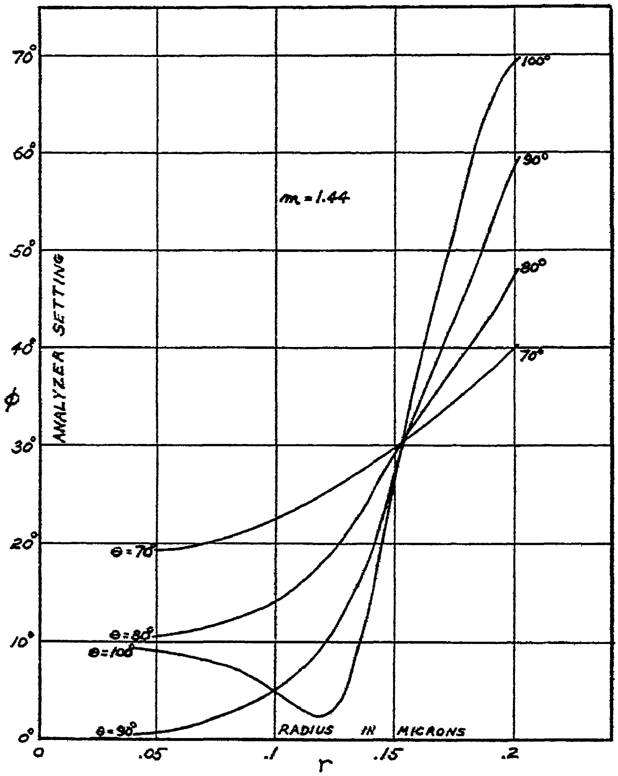


FIG. 10. Calibration curves for the Owl. Analyzer angle ϕ *vs.* r for $m = 1.44$ at four values of θ.

The substance from which the smoke is to be formed is contained in the "boiler", a 2-l. Pyrex flask (see figure 11). The flask and contents are heated electrically in an asbestos board box to between 100°C. and 200°C. depending upon the boiling point of the substance and the particle size desired.

The condensation nuclei are formed in the "ionizer", a 1-l. Pyrex flask fitted with two electrodes sealed into standard-taper joints. The ionizer is mounted above the heater box and connected to the boiler by a standard-taper joint. The condensation nuclei may be formed by a high-voltage electric spark or an electrically heated coil of wire which has been dipped in sodium chloride. Recently (1947) it has been found by Irwin Wilson (24) that the evaporation to dryness of an aerosol produced by atomizing a dilute aqueous solution of sodium chloride in a Vaponefrin nebulizer furnishes an excellent source of radioactive sodium chloride nuclei.

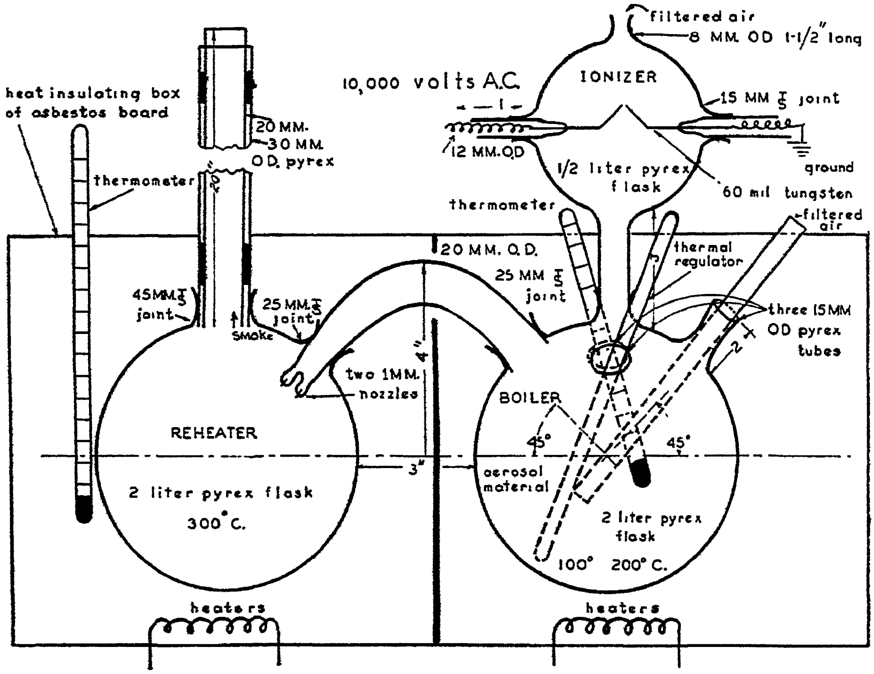


FIG. 11. Generator of aerosols of uniform particle size (improved model)

The "reheater" is a 2-l. Pyrex flask in an adjacent asbestos box heated electrically to about 300°C. A double-walled Pyrex glass chimney 20 in. long is connected to the reheater by a large standard-taper joint. The boiler and reheater are connected by a Pyrex tube having a standard-taper joint at each end. The outlet of this tube into the reheater consists of a jet having two holes of 2 mm. diameter.

A fog is produced by bubbling air through the hot liquid in the boiler, through the glass tube shown in the diagram. At the same time, air is blown in through the ionizer. The total rate of flow is usually from 1 to 4 l. per minute. The mixture of nuclei, spray, and vapor-laden air then passes through the two jet holes into the reheater. Here the spray is vaporized and the nuclei well mixed with the vapor.

The mixture then rises through the chimney, the vapor condensing uniformly upon the condensation nuclei. The aerosol which issues from the chimney is found to be of very uniform particle size. When the column of aerosol is examined in front of a white light it is seen to be brilliantly colored, the colors varying markedly with the angle of observation.

The aerosol has mass concentration of about 1–10 mg. per liter, depending upon the particle size. In order to avoid destroying the uniformity of particle size by coagulation, the aerosol should be diluted immediately 10- or 100-fold with dry, filtered air. Small tubes or jets should not be used for this purpose, since turbulent flow destroys the uniformity of size.

The particle size is increased by increasing the temperature of the boiler, or by increasing the flow of air through the liquid relative to that through the ionizer, or by decreasing the rate of production of ions. A little practice with any given piece of apparatus will show what conditions will yield the most uniform aerosol of a given particle size.

The temperature can be maintained automatically by a thermal regulator inserted directly into the liquid or into the heater box. The air must be dried and well filtered. If it is not, the moisture, dust, and oil fog droplets in the air from a compressor or even from gas stored in a tank will provide so many condensation nuclei that the control of particle size by the ionizer will be lost.

To produce the larger particle sizes, above 1 or 2 microns radius, it is necessary to increase the proportion of vapor by bubbling the air through a porous disc. In later work (1944–45) for the production of DDT–oil aerosols of 10 microns radius for insecticidal purposes, it became necessary to invert the cooling chimney to allow the larger particles to fall out. Care must be taken to avoid decomposition of the material by excessive heating. When the walls become contaminated with decomposition products, a thorough cleaning of the apparatus improves the homogeneity of particle size.

A substance having a range of boiling points or an impurity, particularly of higher vapor pressure, is not suitable for producing a homogeneous aerosol by this method. The different components condense at different rates in the chimney, causing non-uniformity in the particle size. A volatile impurity sometimes condenses so readily that it forms sufficient nuclei to destroy the control of size by the ionizer. Fractional distillation under vacuum to avoid decomposition is often necessary for the production of sufficiently pure materials. Uniform aerosols have been produced from oleic and stearic acids, triphenyl and tricresyl phosphates, rosin, menthol, ammonium chloride, lubricating oil, and Arochlor. The convenient range of particle radius is from 0.05 to 10.0 microns. Smaller sizes may be produced, but it is difficult to measure the size or uniformity.

In general, the larger particle aerosols are relatively more uniform in size. The result arises from the law of growth, which is of the form $r = \sqrt{r_0^2 + kt}$ (24, 25). Here k is a constant depending upon degree of supersaturation and diffusion rate, r_0 is the radius of the nucleus, and r is the radius produced after a time t. For example, with $kt = 0.16$ micron, an initial nucleus of $r_0 = 0.01$ micron yields an r value of 0.40 micron, whereas with a nucleus of $r_0 = 0.1\ \mu$, $r = \sqrt{0.17} = 0.41\ \mu$. With large values of kt relative to r_0^2 any initial variations in r_0

are obliterated. Running the generator with and without condensation nuclei has a very pronounced effect on the uniformity of particle size. The nuclei may be ionized air molecules or molecules of such substances as nitrogen dioxide, hydrogen, oxygen, or ammonia. When a too intense, flaming spark is used, nitrogen dioxide is readily detectable by its odor and color.

REFERENCES

(1) Blumer, Hans: Z. Physik **32**, 119 (1925); **38**, 304, 920 (1926); **39**, 195 (1926).
(2) Bromwich, T. J.: Phil. Trans. **220A**, 189 (1920).
(3) Debye, P.: Ann. Physik **30**, 57 (1909).
(4) Engelhard, H., and Friess, H.: Kolloid-Z. **81**, 129 (1937).
(5) Gucker, F. T., Sr., Pickard, H. B, and O'Konski, C. T.: J. Am. Chem. Soc. **69**, 429 and 2422 (with J. N. Potts) (1947).
(6) Humphreys, W. J.: *Physics of the Air*, p. 554. McGraw-Hill Book Company, Inc., New York (1940).
(7) Jobst, G.: Ann. Physik **76**, 863 (1925).
(8) Johnson, I., and La Mer, V. K.: J. Am. Chem. Soc. **69**, 1184 (1947).
(8a) Kenyon, A. S., and La Mer, V. K.: J. Colloid Sci. **3**, 163 (1949).
(9) La Mer, V. K.: J. Phys. Colloid Chem. **52**, 95 (1948).
(10) La Mer, V. K., and Barnes, M. D.: J. Colloid Sci. **1**, 71, 79 (1946).
(11) La Mer, V. K., Barnes, M. D., Kenyon, A. S., and Zaiser, E. M.: J. Colloid Sci. **2**, 349 (1947).
(12) La Mer, V. K., and Kenyon, A. S.: J. Colloid Sci. **2**, 257 (1947).
(13) La Mer, V. K., and Sinclair, D.: "Portable Optical Instrument for Measurement of Particle Size in Smokes"; "An Improved Homogeneous Aerosol Generator," Report No. 32200, Office of Publications Board, Department of Commerce, and OSRD Report No. 1668.
(14) La Mer, V. K., and Sinclair, D.: "Verification of Mie Theory," OSRD Report No. 1857 and Report No. 944, Office of Publications Board, U. S. Department of Commerce; to appear also as a special publication of the Bureau of Standards Applied Mathematical Computations.
(15) Love, A. E. H.: Proc. London Math. Soc. **30**, 308 (1899); **31**, 489 (1899).
(16) Mie, G.: Ann. Physik **25**, 377 (1908).
(17) Ray, B.: Proc. Indian Assoc. Cultivation Sci. **7**, 1 (1921).
(18) Rayleigh, Lord: *Scientific Papers*, Vol. I, pp. 92–3; Vol. IV, p. 400. Cambridge University Press, London (1899).
(19) Ruedy, R.: Can. J. Research **19A**, 117 (1941); **21A**, 79, 99 (1943); **22A**, 53 (1944).
(20) Sinclair, D.: J. Optical Soc. Am. 37, 475 (1947).
(21) Stratton, J. A.: *Electromagnetic Theory*, p. 563. McGraw-Hill Book Company, Inc., New York (1941).
(22) Stratton, J. A., and Houghton, H. G.: Phys. Rev. **38**, 159 (1931).
(23) van de Hulst, H. C.: *Optics of Spherical Particles*. Duwaer & Sons, Amsterdam, Netherlands (1946).
(24) Wilson, I., and La Mer, V. K.: J. Ind. Hyg. Toxicol. **30**, 265 (Sept. 1948).
(25) Zaiser, E. M., and La Mer, V. K.: J. Colloid Sci. **3**, 571 (December, 1948).

APPENDIX A

Absorbing media

Absorption may arise from two causes: (*1*) in conducting media for which the conductivity, σ, is finite, and (*2*) in dielectrics when the incident wave length is not far from that of an emission line. The first type of absorption is the only one considered in the derivation of the Mie theory. The second type of absorption results from the interaction of bound electrons and the incident electromagnetic wave. The details of the derivation are outside the scope of this paper. As indicated below, this type of absorption may be measured and used in the calculation of light scattering by spherical particles in the same manner as the absorption arising from conduction.

The notion of a complex refractive index is useful whenever absorption occurs. This can be seen by comparing the solutions of the electromagnetic wave equation for transparent and absorbing (conducting) media for which the magnetic permeability is unity. For example, the solution of a plane polarized wave propagated in the $+z$ direction is:

$$E_0 e^{iwt} e^{-ipz}$$

where E_0 is the amplitude of the electric vector, w is the angular frequency, p is the propagation constant, and $i = \sqrt{-1}$. In transparent media p has the value $2\pi m/\lambda$, where λ is the wave length in free space and m is the real refractive index.

In conducting media, owing to the conduction of free electrons, p has the value:

$$\sqrt{\frac{4\pi^2 m^2}{\lambda^2} - \frac{8\pi^2 \sigma}{c\lambda}.}$$

where c is the velocity of light in free space.

By analogy with transparent media the complex quantity, p, when multiplied by $\lambda/2\pi$, yields the complex refractive index:

$$m_c = p\frac{\lambda}{2\pi} = \sqrt{m^2 - \frac{2\sigma}{\nu}i}$$

where m is the real part of the refractive index and ν is the frequency. If now we write:

$$m_c = m(1 - ik)$$

the solution of the wave equation may be expressed in terms which are readily measurable experimentally. Thus:

$$p = \frac{2\pi}{\lambda} m(1 - ik)$$

so that the solution becomes:

$$E_0 e^{-2\pi mk/\lambda} z e^{i[wt-(2\pi mz/\lambda)]}$$

an exponentially damped wave. Since the intensity is proportional to the square of the amplitude, the absorption coefficient, μ, will be given by:

$$\mu = \frac{4\pi k}{\lambda}$$

i.e., the intensity of a plane wave decreases to the fraction $l^{-4\pi kl/\lambda}$ of its original value when the light passes through a distance l of the substance. The quantity,

k, is called the absorption index. It is commonly determined by measurement of the absorption coefficient, μ, in the bulk material at a given wave length. Similarly, the real part of the refractive index, m, must be determined for each wave length, particularly for highly absorbing media, in which m varies rapidly with wave length.

In the case of absorption of wave lengths near an emission line of a dielectric the absorption index, k, and the real refractive index, m, may be measured in the same manner. In either case the complex refractive index $m(1 - ik)$ when substituted in equation 1 or 2 gives the light scattered, i.e., exclusive of that absorbed, by absorbing particles. When $m(1 - ik)$ is used in equation 3, one obtains both the scattered and the absorbed light. Any of the equations 1 to 7 can be used with either a real or a complex refractive index, depending on the nature of the aerosol and the quantity desired.

APPENDIX B

At the request of the referee we give definitions and meanings of a_n and p_n. In the solution of the Maxwell equations, as given by Mie (1908), Debye (1909), and Rayleigh (1911), both the incident wave and the scattered wave are developed into electric and magnetic multipole components. Since the scattered components are coherent, they will interfere. Interference of the scattered wave with the incident wave produces extinction. Mutual interference among the components of the scattered wave determines the intensity of the scattered light. Each electric or magnetic multipole component is characterized by a coefficient. Thus for the 2^n electric component:

$$a_n = \frac{\alpha\psi'_n(\beta)\psi_n(\alpha) - \beta\psi'_n(\alpha)\psi_n(\beta)}{\alpha\psi'_n(\beta)\zeta_n(\alpha) - \beta\zeta'_n(\alpha)\psi_n(\beta)}$$

and for the 2^n magnetic component:

$$p_n = \frac{\alpha\psi_n(\beta)\psi'_n(\alpha) - \beta\psi_n(\alpha)\psi'_n(\beta)}{\alpha\psi_n(\beta)\zeta'_n(\alpha) - \beta\zeta_n(\alpha)\psi'_n(\beta)}$$

Here $\alpha = 2\pi r/\lambda$ and $\beta = m\alpha$.

The cylindrical functions appearing above are defined as

$$\psi_n(\alpha) = \left(\frac{\pi\alpha}{2}\right)^{1/2} J_{n+\frac{1}{2}}(\alpha); \qquad \chi_n(\alpha) = (-1)^n \left(\frac{\pi\alpha}{2}\right)^{1/2} J_{-n-\frac{1}{2}}(\alpha)$$

and

$$\zeta_n(\alpha) = \psi_n(\alpha) + i\chi_n(\alpha)$$

Primes denote derivatives and J represents a Bessel function of half order. In studying the theory the reader will encounter spherical harmonics defined as:

$$\pi_n(v) = \frac{1}{\sin\theta} P_n^1(v) \quad \text{and} \quad \tau_n(v) = \frac{d}{d\theta} P_n^1(v)$$

both derived from $P_n^1(v)$, the first associated Legendre function with $v = \cos\theta$, with the additional useful relation

$$\tau_n(v) = v\pi_n(v) - (1 - v^2)\pi_n^1(v) = \operatorname{cotg} \theta \, P_n^1(v) - P_n^2(v)$$

The functions $\psi_n(\alpha)$ and $\chi_n(\alpha)$ are real when their arguments are real. Both α and β are real if the refractive index m is real, as holds for dielectric spheres. The case of totally reflecting spheres, having m = infinity, may be included.

The cylindric functions behave differently for small and large values of the argument. Thus:

Small α:

$$\psi_n(\alpha) = \frac{\alpha^{n+1}}{1\cdot 3 \cdots (2n+1)}; \quad \chi_n(\alpha) = \frac{1\cdot 3 \cdots (2n+1)}{\alpha^n}$$

represents the first term of a convergent series practicable for $\alpha << n$.

Large α:

$$\psi_n(\alpha) = \sin\left(\alpha - \frac{n\pi}{2}\right); \qquad \chi_n(\alpha) = \cos\left(\alpha - \frac{n\pi}{2}\right)$$

shows the first term of a semiconvergent series practicable when $\alpha >> n$. Debye (1909) has derived semiconvergent developments for intermediate values of α and n, which connect the above formulas, useful when both α and n are large. The computations of the tables by Lowan and associates are based upon the above definitions. The reader should refer to the excellent dissertation of van de Hulst (23) for more details of the meaning of the equations and for improvements in computational schemes, particularly Chapters 2 and 9. In his notation our $\alpha = x$ and our $\beta = y$ (see also Stratton (21)).

APPENDIX C

Definition of i_1 and i_2 in equation 7

The definitions of i_1 and i_2 are obtained from the equations and definitions in appendix B by calculating the energy flux by means of the Poynting vector in the limit for an observer at an infinite distance from the sphere, i.e., $R >> r$.

$$i_1(\theta) = \sum_{n=1}^{\infty} \frac{2n+1}{n(n+1)} \{a_n\pi_n(v) + p_n\tau_n(v)\} \Big|^2$$

$$i_2(\theta) = \sum_{n=1} \frac{2n+1}{n(n+1)} \{p_n\pi_n(v) + a_n\tau_n(v)\}$$

In the van de Hulst reference (23, p. 8), compact formulae are also given for computing the total amount of scattered light and the total forward component by the use of the integrals of products of the functions $\pi_n(v)$ and $\tau_n(v)$ on the sphere as derived originally by Debye.

NUCLEI IN EVAPORATION AND CONDENSATION[1]

W. H. RODEBUSH

Noyes Chemical Laboratory, University of Illinois, Urbana, Illinois

Received October 18, 1948

The condition of dissolved non-polar gases in water solution is entirely a matter of speculation. There is evidence that bubbles small enough to escape observation are very persistent and that the properties of water are affected noticeably by their presence.

The supersaturation of vapors in the absence of nuclei is subject to thermodynamic treatment. The Thomson equation is no longer valid when the single drop ceases to determine a liquid phase. The drop ceases to constitute a liquid phase when the number of molecules becomes small enough so that the share of each molecule in the (colligative) entropy of the drop becomes considerable. The drop now becomes a molecular aggregate in a homogeneous equilibrium and a definite limit to supersaturation is established. Experimental data are in general agreement with the conclusions of this theory.

The nuclei of evaporation and condensation become bubbles or drops. The bubble and the drop are the opposite faces of the coin, so to speak, of which the liquid–vapor interface is the coin disk itself. The low entropy of surface formation of water indicates a high degree of structure in the surface, just as in the body of the liquid. One can hardly avoid the conclusion that water is a unique liquid which always must be discussed as a special case.

That the density of the liquid is greater than the density of ice is a bizarre fact which raises many questions. It seems to be clearly established that ice is a four-coördinated structure with hydrogen bonds maintaining a distance of 2.76 Å. between the oxygen atoms. Since the hydrogen bond is a very weak bond, these linkages can not be considered as fully established except at the lowest temperatures. For the same reason water does not form a glass; the shift to a crystalline structure takes place too readily. As a matter of fact, Pauling considers that ice is only imperfectly crystalline and is to a small degree tautomeric with a glass. If the hydrogen bonds were stronger, one might expect water to form a glass of the same density and coördination number as ice. In liquid water (4) the average distance between oxygen atoms is increased to about 3.0 Å. The increased density is accounted for by the fact that the coördination number is on the average greater than 4.0, perhaps as much as 4.5. Since the hydrogen bonds are constantly breaking and re-forming, this is only an average figure. The increased density is easily understood, even though the average distance between molecules is greater. Likewise the heat of fusion is easily accounted for by the fact that both the hydrogen-bond energy and the van der

---

[1] Presented at the Symposium on Aerosols, which was held under the joint auspices of the Division of Physical and Inorganic Chemistry and the Division of Colloid Chemistry at the 113th National Meeting of the American Chemical Society, Chicago, Illinois, April 22, 1948.

Waals energy will be considerably decreased by an increase in interatomic distances.

Not all of the properties of liquid water are so easily accounted for, however. For example, consider the solubility of gases, such as oxygen or methane, which are a long way above their critical temperatures. It is easily understood that, although ice has an open structure, the small tolerances of the lattice do not allow the inclusion of foreign molecules. How then can the liquid phase, which is denser, make room for the dissolved gas?

If one assumes that the gas molecules are simply entrapped in the mesh of the liquid water, then the solubility should increase with temperature. It does in some cases but usually not. For example, the heat of solution of methane is negative.

Finally there is the strange group of hydrates formed by the gases such as chlorine, methane, and argon. These hydrates are believed to contain eight molecules of water per molecule of gas. Since the unit of structure in hexagonal ice is a group of eight water molecules, it has been suggested by some that the hydrates are made up by enclosing one molecule of gas in one of the ice units, but this idea is untenable because there is no room for a gas molecule in such a cluster.

FOREIGN NUCLEI

It might be supposed that the theory of fluctuations would account for the initiation of all types of processes in the liquid state in which a new phase appears, but it is an interesting fact that most of the investigators who have studied processes of this type in water solution have been led to postulate the presence of foreign substances in some form of dispersion or solution. Among the phenomena which have been studied are the coagulation of sols (3), the tensile strength and superheating of water (2; see also 6), and the effect of centrifuging on the E.M.F. of a galvanic cell. One may conclude in every case that they are due to the presence of impurities and note that treatment of the water by such drastic processes as ultrafiltration, ultrasonic radiation, or high pressure often causes these effects to disappear. Without generalizing in any way as to how these effects are produced, one may assume that in some cases at least the homogeneous structure of water vapor is disturbed by the presence of dissolved gases in various states of dispersion.

We shall consider, therefore, what the behavior of a dissolved gas such as oxygen might be. If the gas is not molecularly dispersed, the bubbles to escape detection must be less than 10^{-7} cm. in radius. Such a bubble at 1 atm. would not contain more than one molecule, but at the high internal pressure existing in water the density might well be greater than that of liquid oxygen at the critical temperature. If one wonders how a bubble at such a pressure can exist in equilibrium with the external atmospheric pressure, it is only necessary to remember that this pressure is largely an internal pressure due to forces acting throughout the liquid. The only pressure acting on the bubble tending to cause it to disappear is the pressure due to interfacial tension. The interfacial tension between oxygen and water on a flat surface is large. For limited surfaces of high convexity it might be small. The very large partial molal volume of these gases in water solution lends some support to the hypothesis of a low interfacial tension at the surface of small bubbles. If the bubble were under a very high pressure, the partial molal volume could scarcely be so great.

It is not surprising that the solubility of oxygen in water is small, but it is paradoxical that the temperature coefficient of solubility is negative. This implies that the heat of solution is negative; in other words, that the attraction of water molecules for oxygen molecules is greater than the mutual attraction between water molecules. But if this is the case, the solubility should be greater in the first place.

We shall have to try for another explanation. The small solubility means a negative entropy of solution. This negative entropy of solution could be due to a very definite orientation of water molecules in the interface. The energy liberated by this orientation would account for the negative heat of solution without assuming strong attraction between oxygen and water molecules.

This resembles the "iceberg" theory (1), so-called, but there can be no resemblance to an ice structure. The ease with which water supercools argues against any structure resembling hexagonal ice forming in the liquid state. The reason for this is obvious of course: the high internal pressure makes the ice structure unstable because of its low bulk density. A pentagonal dodecahedron would form a somewhat larger unit (thirty molecules) than the hexagonal unit of eight molecules and there would be room on the inside for a gas molecule. It seems probable, however, that the most stable configuration would be a bubble containing a number of gas molecules. The water molecules in the surface of this bubble would be oriented in such a way as to form at least three hydrogen bonds per molecule. The fourth hydrogen would project on the inner surface of the bubble in such a way as to be attracted strongly to gas molecules by polarization forces. This arrangement would give a low or zero interfacial tension and a negative heat and entropy of solution.

There will be a size of bubble for which the heat of solution and entropy will be a maximum (negative) but the distribution will be statistical. It may be that equilibrium can be established only at the gas–liquid surface, so that the distribution of size may be far from that which would be predicted. There may be more bubbles containing only one molecule than of any other size, but these may represent only a small percentage of the total gas dissolved.

CONDENSATION OF WATER VAPOR

In discussing the condensation of vapor we shall treat only the case in which foreign nuclei such as ions and dust particles are assumed to be absent. In the vapor of a hydrocarbon oil there may be a few large molecules which serve as nuclei for condensation. In water vapor in a pure state the nuclei must be aggregates of water molecules which grow by accretion, starting with the association of two water molecules to form a bimolecular complex. We shall refer to these aggregates as clusters until they become large enough to form drops. We shall define the terms "cluster" and "drop" somewhat arbitrarily by applying a distinguishing criterion of thermodynamics. If the aggregate contains enough molecules so that its behavior in the presence of vapor at a given temperature and pressure is uniquely determined, independently of the number of such aggregates present in a given volume, then it constitutes a liquid phase and we shall call it a drop. On the other hand, if the aggregate contains a small number of molecules so that its behavior is statistical and depends upon the concentration, we shall call it a cluster. There is of course a gradual transition from cluster to drop which must occur somewhere between 100 and 1000 molecules per aggregate, i.e., in the range of radius 10^{-7} to 2×10^{-7}. As we shall see later, a drop will determine an equilibrium saturation pressure and hence can no longer be considered as a nucleus.

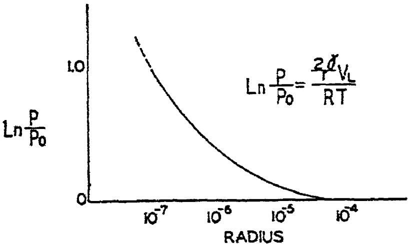


FIG. 1 Plot of Thomson equation

THE THOMSON EQUATION

The Thomson equation

$$\ln \frac{p}{p_0} = \frac{\frac{2\gamma}{rV}}{RT}$$

cannot be assumed to hold below a radius of 10^{-7} (133 molecules) for two reasons: (*1*) As the number of molecules decreases, the concept of a flat surface under uniform tension must fail. (*2*) It is obviously impossible to speak of a saturation pressure when the equilibrium pressure depends upon the concentration of clusters present, as it obviously does in the limit when only bimolecular clusters are present. So long as the saturation pressure is fixed by the presence of a single drop we are dealing with vapor–liquid equilibrium, but when the aggregates become small enough so that their concentration is significant in determining equilibrium we have a homogeneous instead of a heterogeneous equilibrium. Accordingly we have drawn the curve for $\ln p/p_0$ (figure 1) down to $r = 10^{-7}$ ($\ln p/p_0 = 1$). The curve is extended beyond the point as a broken line. This broken line indicates not only that the Thomson equation cannot hold but that *no saturation pressure exists for this region* ($< 10^{-7}$).

THERMODYNAMIC EQUATIONS

The equilibrium between vapor and liquid drops or molecular clusters may be conveniently expressed by the equation:

$$S_v - S_L = \frac{\Delta H_V}{T}$$

Here S_v, the entropy per mole of vapor, is given by the equation

$$S_v = S^0 - R \ln P$$

where S^0 is the standard entropy at unit pressure. S_c, the entropy per mole of the liquid drop of molecular cluster, is represented by

$$S_c = S_L + \frac{1}{n}(S^0)_{col} - \frac{2\gamma V}{T}$$

Here S_L is the entropy per mole of liquid in the mass; $(S^0)_{col}$ is the contribution of translational, rotational, and concentration terms to the entropy of aggregates composed of n molecules each and treated as a gaseous phase. The effect of surface tension in reducing the entropy of the liquid comes directly out of the Thomson equation, and we shall assume ΔH_v, the heat of vaporization, to be constant so long as the Thomson equation holds. The entropy curves for liquid and vapor are plotted in figure 2 as a function of radius.

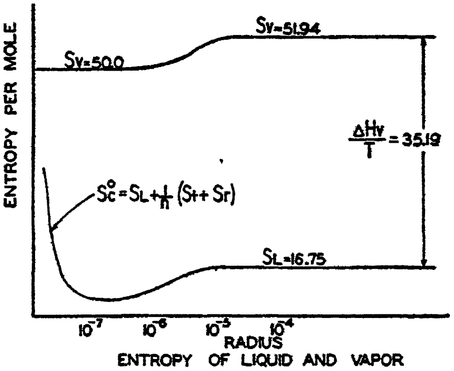


FIG. 2. Plot of entropy curves for liquid and vapor as a function of radius

THE MINIMUM ENTROPY FOR SMALL DROPS

The entropy curve for the liquid shows a minimum in the neighborhood of radius 10^{-7}. This minimum marks the transition from cluster to drop, in other words, the limit of the saturation curve corresponding to a liquid–vapor equilibrium. The existence and approximate location of this minimum require justification, and this justification can be given both from theoretical considerations and from the behavior of supersaturated vapor as reported by investigations.

The maximum decrease in entropy due to surface tension is about two units. Beyond this point a new effect becomes important, *viz.*, the colligative entropy. Because the aggregates representing the liquid phase are in suspension, they may be treated as large molecules forming a part of the homogeneous vapor phase and possessing translational, rotational, and volume concentration entropy. The share of this entropy per molecule is only $1/n$ times the total, where n is the number of water molecules per aggregate. When the number of molecules per aggregate is less than 100, this contribution will more than offset the decrease in entropy due to surface tension and hence a minimum in the entropy of the liquid phase must occur at or near this radius (10^{-7}).

One can calculate the entropy of the drop and estimate the entropy of the bimolecular cluster. (The heat of dissociation of the bimolecular cluster is estimated at 5000 cal. and entropy of dissociation at 20 E.S.U.) The entropy of cluster of intermediate size is obtained by an interpolation made on the assumption that the decrease in heat of vaporization is offset by an increase in vibrational entropy. This last assumption is a variation of the Trouton rule and quite reliable. Hence one locates the minimum entropy as occurring at or near a radius of 10^{-7}.

This estimate based on theoretical calculations receives a very satisfactory confirmation from the experimentally reported behavior. It is, of course, not easy to reconcile all of the facts as reported in the literature.[2] One cannot be sure of the absence of foreign nuclei in all cases. Supersaturation in cloud chambers is complicated by the presence of air and condensation on the walls. The most reliable experiments are those in which superheated steam is expanded through a nozzle.

The following conclusions seem to be established on a fairly reliable basis for pure water vapor: (*1*) A threefold saturation ($\ln p/p_0 \sim 1$) does not result in rapid condensation. (*2*) If the saturation is increased to four- or fivefold ($\ln p/p_0 = 1.5$) condensation will take place in 20–40 microsec. with a drop concentration of 10^6–10^8 per cubic centimeter.

THE SUPERSATURATION REGION

It does not affect the conclusions of this paper seriously if the above limits are not strictly accurate. One can be certain that certain limits of supersaturation exist such that at the lower limit condensation does not occur and at the upper limit it occurs very rapidly. These limits are separated by a difference in the pressure of the vapor that corresponds to about one entropy unit. In figure 3 the equilibrium concentrations of clusters are shown for the lower and upper limiting pressures. The ratio of the concentrations between the two curves is given by the relation

$$R \ln C = n\Delta S$$

where ΔS is taken as one entropy unit.

THE BOTTLENECK FOR GROWTH OF NUCLEI

The lower curve in figure 3 is for $\ln p/p_0 = 1$ and shows a minimum in the neighborhood of $r = 10^{-7}$ corresponding to the minimum entropy for the condensing phase. At this minimum the probable concentration of aggregates is vanishingly small. Since condensation must take place by the growth of clusters with drops, and since this growth cannot take place unless a large concentration of clusters is present at equilibrium, no condensation will occur. If the probability of the existence of a cluster is small, the cluster will not grow but will diminish in size.

---

[2] Ruedy (5) has summarized the experimental data on supersaturation in the literature.

If, however, the supersaturation is increased to $\ln p/p_0 = 1.5$, we have the upper curve with a minimum somewhat below $r = 10^{-7}$ but at concentrations high enough to permit a rapid growth of clusters into drops. There is no point in trying to draw the concentration curve beyond the minimum, because it represents a concentration that will never be realized since supersaturation of the vapor is no longer possible in this region.

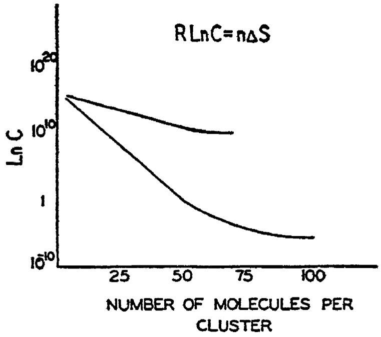


FIG. 3. Equilibrium concentrations of clusters for the lower and upper limiting pressures

MECHANISM OF GROWTH OF NUCLEI

Bimolecular complexes must be formed by triple collisions. There are some 10^{24} triple collisions per second at the threefold saturation; hence a concentration of 10^{16} bimolecular complexes can be formed in a matter of microseconds. A cluster grows by individual accretion. The collisions are inelastic but, of course, for each molecule condensing another must evaporate, otherwise the temperature of the cluster would rise without limit. Actually, the temperature of the cluster will rise until the rate of evaporation is equal to the rate of condensation from the supersaturated vapor. For a threefold saturation this would amount to 15 or 20°. About 150 cal. per mole will be lost on each exchange of molecules, so that perhaps one molecule will adhere permanently for each hundred exchanged. The rate of growth may be calculated by the kinetic theory to be $4.65 \times 10^{21} \times r^2$ for threefold saturation at 25°C. A cluster would grow to a radius of 10^{-7} cm. in a few microseconds.

GROWTH BY AGGLOMERATION

If the clusters are small so that the Brownian motion is rapid and the concentration high ($> 10^{10}$), then collisions between clusters will be numerous and will result in a rapid growth by agglomeration with a corresponding reduction in the total number present. This accounts for the fact that in a rapid condensation the number of droplets per cubic centimeter will not exceed 10^8. However, if the clusters are small the heat liberated is so great that many molecules must be evaporated and little gain in size will result from the collision. The reason for this is that the heat of condensation increases rapidly with increasing size of the cluster and levels off when it becomes a drop of 10^{-7} radius. For example, let us suppose the heat of condensation per mole (including surface energy) in a cluster of 50 molecules to be 5000 cal. and in a cluster of 100 molecules to be 7500 cal. The heat of agglomeration would be so great as to require the evaporation of nearly one-third of the molecules contained in the two drops.

In this paper no attempt has been made to calculate the over-all rate of growth of clusters of sizes and hence the absolute rate of condensation. Such a calculation undoubtedly is possible but laborious.

It is not immediately obvious that the considerations of this paper apply to the condensation of water vapor in the atmosphere because foreign nuclei are usually present and hence the supersaturation is small. It is conceivable, however, that under extreme conditions a high degree of supersaturation may be reached.

REFERENCES

(1) Frank, H. S., and Evans, M. W.: J. Chem. Phys. **13**, 507 (1945).
(2) Harvey, E. N., *et al.*: J. Am. Chem. Soc. **67**, 156 (1945).
(3) La Mer, V. K.: J. Phys. Colloid Chem. **52**, 65 (1948).
(4) Morgan, J., and Warren, B. E.: J. Chem. Phys. **6**, 666 (1938).
(5) Ruedy, R.: Can. J. Research **A22**, 77 (1944).
(6) Scott, A. F., Shoemaker, D. P., and Wendell, J. G.: J. Chem. Phys. **16**, 495 (1948).

NUCLEATION OF SUPERCOOLED WATER CLOUDS BY SILVER IODIDE SMOKES[1]

BERNARD VONNEGUT

General Electric Research Laboratory, Schenectady, New York

Received October 18, 1948

ABSTRACT

Measurements have been carried out in a cold box on the nucleation of ice formation in supercooled clouds by silver iodide smoke particles. The smokes were produced by atomizing an acetone solution of silver and ammonium iodides into a hydrogen flame.

The measurements show that ice crystals do not form immediately on the silver iodide particles. The silver iodide can be regarded as greatly increasing the probability of ice crystal formation. The number of effective nuclei which can be produced per gram of silver iodide varies from about 10^{13} at −10°C. to 10^{16} at −20°C.

INTRODUCTION

About a year and a half ago, it was found that silver iodide smokes had the property of causing snowflakes to form in a supercooled cloud (5). It is believed that silver iodide particles are good nuclei for ice formation because of the close resemblance of their crystal structure to that of ice. Experimentation has been under way to learn more about the production and behavior of these smokes. The work to be described in this account should be regarded as preliminary. The techniques and apparatus used in the work frequently leave much to be desired in the way of precision. The results are tentative and await confirmation by better experiments. This work, despite the uncertainties in it, nevertheless sheds light on the mechanism of nucleation and suggests new experiments and improved techniques for giving a more complete picture of the phenomena associated with nucleation by silver iodide.

APPARATUS AND TECHNIQUES FOR MEASUREMENTS ON SMOKES

Wind tunnel

In order to determine the output of a source of silver iodide smoke, it is necessary to secure a sample of smoke for testing that is a known fraction of the total output. This was accomplished by diluting the output of the smoke generator with a large known flow of air, and taking a known volume of this dilute smoke for testing. The smoke generator was placed in front of a $\frac{1}{4}$-horsepower electric fan, 3 ft. in diameter, which sucked the smoke along with a large volume of air into a crude wind tunnel 4 ft. square in cross-section and 24 ft. long. The stirring action of the fan and the turbulence in the tunnel mixed the smoke with the air to produce a more dilute smoke, which was discharged at the other end of the tunnel where samples were taken. When a heavy white oil smoke was introduced into the tunnel, it appeared to be quite uniformly mixed and diluted as it left the tunnel. The rate of flow in the tunnel, as determined by measuring the velocity with a vane-type anemometer, was 4 × 10^6 $cm.^3$ per second.

---

[1] Presented at the Symposium on Aerosols which was held under the joint auspices of the Division of Physical and Inorganic Chemistry and the Division of Colloid Chemistry at the 113th National Meeting of the American Chemical Society, Chicago, Illinois, April 22, 1948.

Of the silver iodide smokes used, all except the ones having the largest particle size were completely invisible under the conditions of the experiment. The smokes having the largest particles were quite transparent and of a pale blue or purple appearance in the sunlight.

Smoke sampling and diluting syringe

In many of the cases, only a fraction of 1 cc. of the smoke from the tunnel is needed for a test. Precipitation on the walls of a container of this size would be very rapid. Therefore, a syringe was constructed for taking a sample of the smoke and diluting it quantitatively to any desired amount. The syringe (figure 1) consists of a metal tube 3 in. in diameter with a piston and leather washers which can be moved back and forth a fixed distance. A sample of smoke was taken with the syringe and then diluted to the desired concentration by moving the piston in and out the requisite number of times in smoke-free air, upwind from the tunnel.

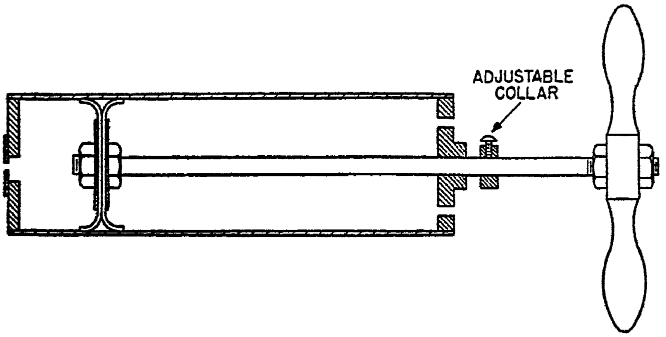


FIG. 1. Smoke sampling and diluting syringe

Cold chamber

The early measurements on the number of nuclei contained in silver iodide smokes were made using Schaefer's technique (3). A measured volume of smoke was introduced into a supercooled cloud in a home freezer and the number of snow crystals produced per cubic centimeter was visually estimated.

This technique has been slightly modified in the more recent work and the apparatus used is shown in figure 2. The tests were carried out in a brass cylinder 15 in. high and 12 in. in diameter having walls ½ in. thick to provide good thermal conductivity.

This chamber, which was closed at the bottom, was maintained at a low temperature by placing it in a 4 cu. ft. home freezer. The freezer thermostat was used to regulate the temperatures. A supercooled cloud was maintained in the refrigerated cylinder by evaporating water from wet paper toweling wound around a 15-watt electric heater placed in the lower part of the cylinder. A hinged masonite lid was used to close the top of the cylinder during tests. The temperature at the top of the cylinder was found to be about 2°C. warmer than at the bottom. The minimum temperature obtainable in the chamber was −20°C. In future experiments, it is highly desirable that lower temperatures be obtainable and that provisions be made for better temperature regulation.

Smokes were tested by introducing them from the sampling syringe into the supercooled cloud in the cylinder. A stack of cold microscope slides was placed in the bottom of the cylinder. The snowflakes produced by the action of the smoke settled on the bottom of the cylinder and on the topmost microscope slide. At intervals 2 min. apart the slide on which the snow had fallen was removed, thus exposing the slide beneath. The slide which was removed was then examined under a microscope kept in the freezer. By means of a Whipple eyepiece, the number of flakes collected per square millimeter in a 2-min. period was counted. When the rate of snowfall had dropped to a low value, the total number of flakes which had fallen per square millimeter was determined by adding the numbers which had fallen in each 2-min. sample. The total number of snow crystals which would have been produced by the entire output of the generator could then be calculated from the area of the bottom of the cylinder, the volume and dilution of the smoke introduced, and the volume rate of production of the smoke.

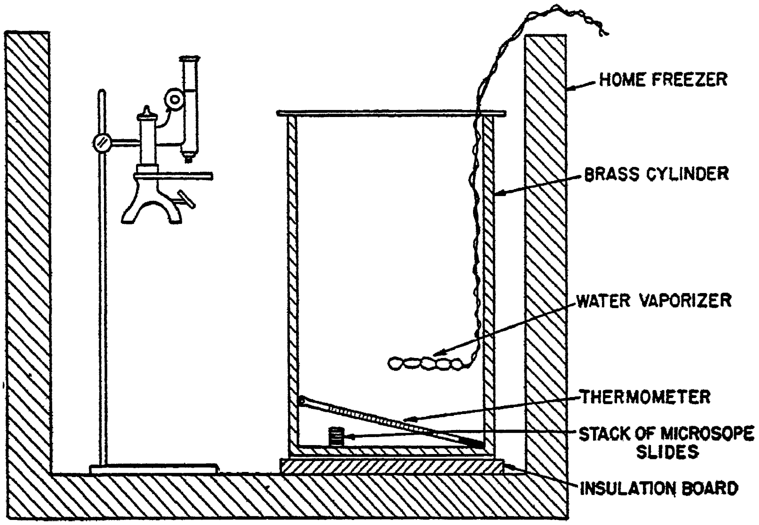


FIG. 2. Apparatus for counting nuclei in smoke

Electron microscope examination of smoke

Smokes being tested were examined with the electron microscope to determine their appearance and particle size. Samples of smoke were precipitated on a Formvar film supported on a fine wire screen, by moving it for about a minute in and out of the smoke stream about 3 ft. from the generator. The smoke stream at this point is still quite warm. The sample screen, because it is in the smoke only a moment at a time, remains cool so that thermal precipitation may play a part in the collection of the smoke.

Under these conditions of precipitation, it is quite likely that particles of a certain size may be selectively precipitated, so that the sample obtained is not entirely representative of the smoke. A more reliable method would be desirable so that more trustworthy data could be obtained.

The smoke samples were photographed using the electron microscope. Determinations of the particle size and the number of particles per cubic centimeter of material were made from these photographs.

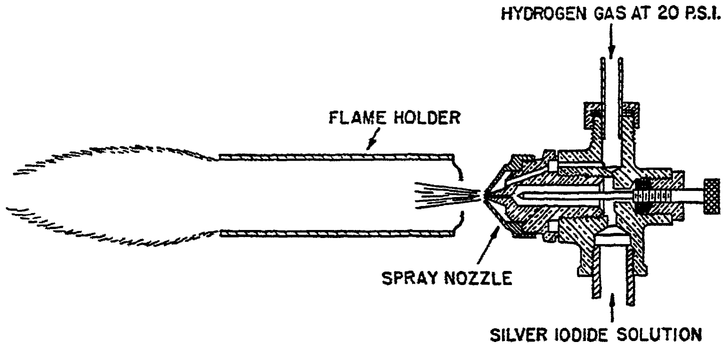


FIG. 3. Silver iodide smoke generator

Smoke generator

The smokes used in these tests were produced by a smoke generator constructed from a commercial compressed-air atomizing nozzle of the sort used for paint spraying and humidifying (Binks No. 174). The generator is shown in figure 3. Compressed hydrogen gas at 20 lb. per square inch was applied instead of air to the air inlet of the nozzle and a solution of silver iodide was used as the liquid to be sprayed. The hydrogen stream as it left the nozzle was ignited. The heat of the flame vaporized the silver iodide in the spray into a gas which, upon mixing with the atmosphere, condensed into a smoke of small silver iodide particles. In order to prevent the hydrogen flame from being blown out by the wind, a flame holder consisting of a piece of ¾-in. pipe 2½ in. long was placed ½ in. from the spray nozzle. At a pressure of 20 lb. per square inch, the spray nozzle used about 3 cu. ft. of hydrogen per minute (measured at atmospheric pressure).

Silver iodide solutions

Although silver iodide is very insoluble in water and organic liquids, it is quite soluble in acetone or water solutions containing a soluble iodide such as sodium or ammonium iodide. The solutions used in this work were made by dissolving 200 g. of silver iodide and 100 g. of ammonium iodide in a mixture of

750 cc. of acetone and 250 cc. of water. A dilute solution was also used which was made by diluting the above solution to ten times its volume with acetone. Solutions can be diluted any desired amount with acetone; however, dilution with water causes precipitation of the silver iodide. Ammonium iodide was used in the solutions for these experiments because it probably is completely decomposed in the hydrogen flame, thus leaving a smoke of uncontaminated silver iodide.

The rate at which silver iodide was fed into the smoke generator was varied by controlling the rate of flow of solution by adjusting the valve on the nozzle and by using solutions of different concentrations.

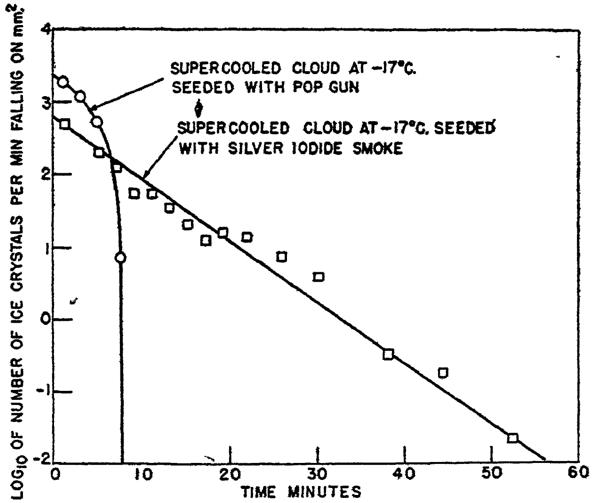


FIG. 4. Comparison of seeding with silver iodide smoke and seeding with popgun

EXPERIMENTAL RESULTS

Decrease in rate of snow formation after seeding

There was found to be a large difference in the behavior of the supercooled cloud when it was seeded with the low-temperature air produced by a popgun and when it was seeded with silver iodide smoke. From figure 4 it can be seen that although many snow crystals were produced by the low temperature from the popgun, all of these crystals had precipitated to the bottom of the container at the end of 10 min. When the cloud was seeded with silver iodide smoke, however, a measurable number of ice crystals were still precipitating at the end of almost an hour. The rate of snowfall decreases to one-half each 2 or 3 min. This rate of decrease was not found to vary significantly with temperature or with the particle size of the smoke although, as will be seen, the total number of snow crystals varied over several factors of ten, depending on the temperature of the supercooled cloud.

Precipitation of smoke in syringe

One possible source of error in these experiments is that which might be caused by coagulation and precipitation of the smoke in the sampling syringe. In order to evaluate this rate of disappearance, tests were made in which small smoke samples were withdrawn from the syringe after it had been in the syringe for varying periods of time. The number of effective nuclei in a sample of smoke was found to decrease by one-half every 20 min. The time required to take, dilute, and discharge a sample into the cold chamber was never more than a minute or two, so that the changes in the smoke occurring during this time were not large.

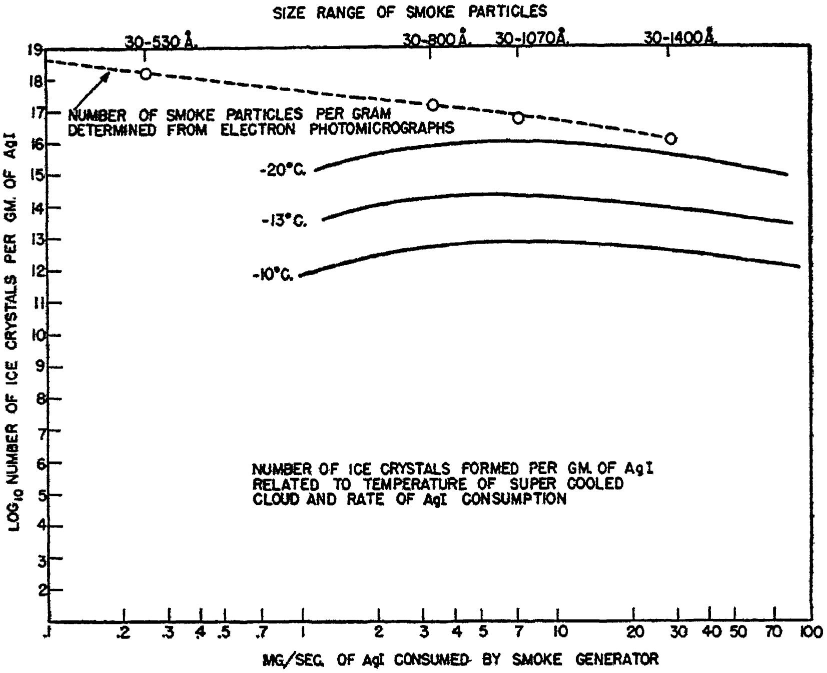


FIG. 5. Number of ice crystals formed per gram of silver iodide related to temperatuer of supercooled cloud and rate of silver iodide consumption.

Number of ice crystals per gram of silver iodide

The results of experiments carried out with different settings of the smoke generator and in supercooled clouds at different temperatures are shown in figure 5. In this graph, the results are also given for the electron microscope examination.

Electron photomicrographs were made of four different samples of smokes. A hundred or more particles in a representative section of each photograph were measured, and a graph was made of the number of particles as a function of their diameter. All of the smokes examined showed a large number of particles having a diameter of about 30 Å., which is the smallest particle observable with the microscope. It, therefore, seems probable that in these smokes there are many more particles which were too small to be resolved. By neglecting these invisible particles, it was possible to estimate the minimum number of particles formed per gram of silver iodide by dividing the number of particles in a representative section of a photograph by the sum of their masses computed on the assumption that they were spheres having the same density as silver iodide crystals. Figure 6 shows electron photomicrographs of two typical smokes.

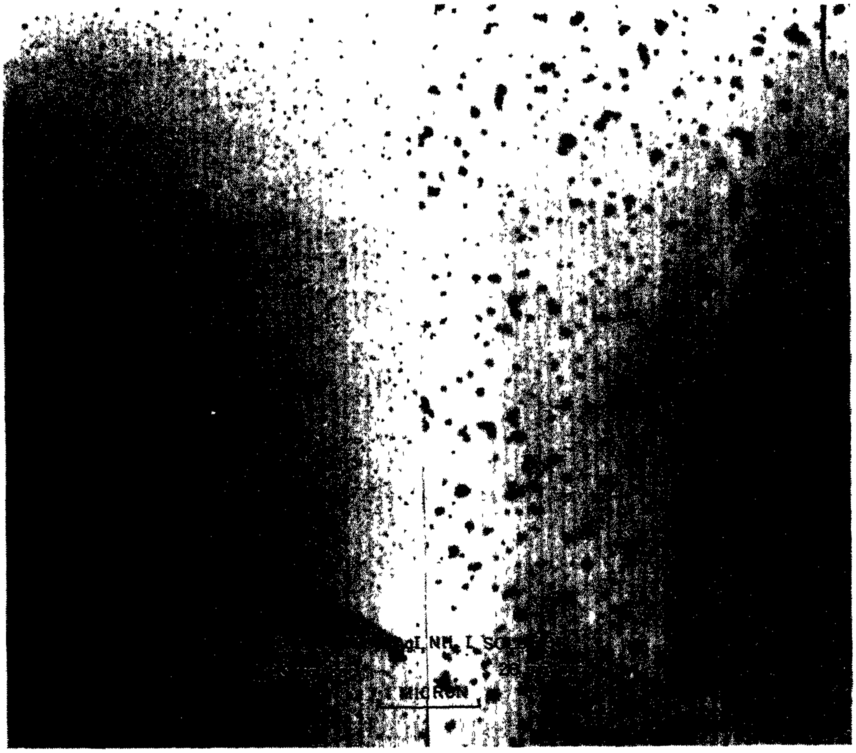


FIG. 6. Smoke from silver iodide–ammonium iodide solution. (a) 0.25 mg./sec.; (b) 28 mg./sec.

DISCUSSION OF RESULTS

The work thus far is not of sufficient scope to give straightforward answers to many questions which arise concerning the mechanism of silver iodide nucleation. It is the purpose of this discussion to venture some possible explanations for the experimental observations.

Formation of ice crystals

Figure 4, which shows the number of snow crystals precipitating per minute after seeding, illustrates a significant difference between the behavior of a cloud seeded with silver iodide and one seeded by a popgun. Seeding with a popgun or dry ice produces very large numbers of small ice crystals by cooling a small region of the cloud to a temperature at which spontaneous nucleation takes place. These ice crystals then mix with the cloud and grow at the expense of the supercooled water drops. From figure 4 it can be seen that although a large number of ice crystals is produced by the popgun, all of these crystals have precipitated out at the end of 10 min. Now if all of the silver iodide particles put into the cloud formed ice crystals at the time they were introduced, they, too, should have precipitated out at the end of 10 min. However, at the end of almost an hour, ice crystals are still precipitating. This leads one to the conclusion that all of the silver iodide particles do not form ice crystals immediately and that even at the end of half an hour or more, appreciable numbers of silver iodide particles have not yet formed snow crystals and are still present in the cloud.

There are various possible explanations for the time required for silver iodide particles to initiate the formation of ice crystals in a supercooled cloud. One possibility is that in order for an ice crystal to form on a crystal of silver iodide, a certain critical number of water molecules must by chance arrange themselves on its surface in the structure of ice. According to this explanation, a silver iodide particle would not act as a nucleus until this event took place. The presence of a silver iodide surface might be regarded as merely greatly increasing the probability of the formation of ice.

If we take a simplified view of the theory of nucleation as advanced by Gibbs, in supercooled water or in a region supersaturated with respect to ice, the water molecules by chance, from time to time, arrange themselves in the lattice of crystalline ice. If these minute aggregations are smaller than a certain size, their vapor pressure is greater than that of the supersaturated region and they are unstable and break up. If, on the other hand, they are larger than a certain size, their vapor pressure is such that they are stable and continue to grow. The lower the temperature, the smaller will be the critical size necessary for stability. Schaefer (4) found, using clean air free of dust, that the rate at which nuclei form is very low at temperatures above −38.9°C. However, in the presence of a silver iodide surface, it is possible, because of the close similarity between ice and silver iodide, that the probability of the chance formation of a nucleus is greatly increased and is large even at temperatures as high as −10°C.

Another way of looking at the phenomenon is to consider the growth of an ice crystal as the formation of a crystal of ice on ice. We know that this takes place with the greatest ease. The formation of ice on a large surface of ice requires the formation of very little new ice surface; hence there is little change in surface energy. For every new ice surface formed, an almost equal ice surface is covered up. However, when a new ice surface is formed in the absence of any other surface, large amounts of surface energy are required relative to the free energy decrease in the formation of the interior of the new phase. Because of the close similarity of ice and silver iodide, the formation of ice on a silver iodide surface probably involves only a small amount of surface energy, and therefore the chances of this taking place are good.

The lower the temperature, the smaller will be the critical size of a stable nucleus. Because the critical size is small at low temperatures, the chances of a nucleus forming are, in general, better. Hence, the rate of nucleus formation is generally greater at low temperatures.

Another possible explanation for the time required for silver iodide particles to form ice crystals can be based on the assumption that the silver iodide particles produced by the generator are not all in the hexagonal form which is similar to ice. Possibly they might exist in a metastable condition as a supercooled liquid or as some other modification. In this case, the rate at which they react to form ice crystals would be determined by the rate at which they transform into the stable hexagonal form. However, it appears for several reasons that this factor is probably not large in these experiments. It has been found that the silver iodide smokes used in these tests will nucleate silver iodide solutions supersaturated with respect to the hexagonal form and that the crystals which result are of the hexagonal variety. The number of such nuclei in a given volume of smoke is of the same order of magnitude as the number of particles determined by the electron microscope. It, therefore, seems probable that at room temperature most of the smoke particles are of the hexagonal structure. In addition one would expect, if the rate-determining factor involved the transformation of the silver iodide particles themselves, that smokes aged for a period of time at low temperature would show a different rate of falling off of ice crystal formation when they are put in a supercooled cloud. This has not been observed.

Another possibility is that the formation of an ice crystal under the influence of a silver iodide particle can take place only in the liquid phase. In this case, the rate of ice crystal formation might be limited by the rate at which silver iodide particles entered into supercooled water drops. This could occur either by diffusion of the particle to the drop or by condensation of a drop on a particle. It has been found that when a suspension of silver iodide particles in water is sprayed into a supercooled cloud, ice crystals are formed. It is therefore probable that silver iodide particles can act as nucleating agents in a drop of liquid water. However, it has not been definitely established that this is a necessary condition for silver iodide to act. If we assume this to be the rate-governing factor in these experiments, we are faced with the problem of explaining the large effect of temperature on the total number of ice crystals produced. As will be discussed later, the variation of the number of ice crystals produced with temperature can be explained on the basis that the rate of nucleation increases greatly with decreasing temperatures. There is no reason to expect that the rate at which particles enter water drops is particularly temperature-sensitive. It is quite possible that the time required for a particle to enter a water drop is of considerable importance in these experiments, but it is probably not the most important rate-governing factor.

The interpretation of results is further complicated by the possibility that small particles of silver iodide smoke may dissolve in water drops before they have a chance to start the formation of an ice crystal. At 25°C. the solubility of silver iodide in water is about 1×10^{-8} moles per liter (1). A water drop 10 microns in diameter on this basis should be capable of dissolving a silver iodide particle about 100 Å. in diameter. The solubility should be somewhat greater for particles of this small size. However, the effect of particle size should be counteracted to some extent by a decreased solubility at the lower temperatures. If these factors are of importance, the kinetics of the nucleation may be dependent to a large extent on the rate of solution of the silver iodide particles in the water drops.

The author is inclined to favor the explanation that the rate of ice crystal formation is governed primarily by the rate of spontaneous ice nucleus formation on the silver iodide particles. In some as yet unreported experiments on the nucleation of supercooled water and supercooled tin, the author measured at constant temperature the rate of solidification of systems composed of a number of independent supercooled drops. It was observed that the drops did not all freeze at once but froze at a rate which steadily decreased with time. This rate of freezing increased greatly as the temperature was lowered. In these experiments, foreign particles and surfaces were present which undoubtedly served as foreign nuclei. The time required for these drops to freeze could be best explained on the basis of the chance formation of stable nuclei on the foreign surfaces. It seems reasonable to believe that in the case of nucleation by silver iodide similar phenomena play a dominant rôle.

If we proceed on the assumption that silver iodide merely increases the rate of the chance formation of spontaneous ice crystals, it is possible to come to some conclusions as to the magnitude of the rate and its dependence on temperature. Some of the particles of silver iodide introduced into the cold chamber undoubtedly become lost as potential nuclei by precipitation on the walls of the chamber. The rate of precipitation on the walls is probably large because of thermal diffusion resulting from the 15-watt heater in the vaporizer. One would expect the rate at which particles disappear by precipitation on the walls to be proportional to the concentration of the particles. On the basis of this assumption, the concentration of particles would decrease exponentially with time. If the rate of snow formation is proportional to the concentration, this too should decrease exponentially with time. This is experimentally observed to be the case.

On the basis of the foregoing assumptions, one can analyze the situation mathematically.

If c is the number of silver iodide particles per unit volume at a time t then:

$$\frac{dc}{dt} = -(K_1 + K_2)c \tag{1}$$

where K_1 is the rate at which the silver iodide particles form ice crystals and K_2 is the rate at which they are being removed from the cloud by precipitation or other causes. By integrating equation 1 we obtain an expression for the concentration c at any time:

$$c = c_0E^{-(K_1+K_2)t} \tag{2}$$

where c_0 is the concentration at time $t = 0$. If N is the number of ice crystals formed per unit volume, then

$$\frac{dN}{dt} = K_1 \cdot c = K_1 c_0 E^{-(K_1+K_2)t} \tag{3}$$

which when integrated gives for the number of crystals formed at time t:

$$N = \frac{K_1 c_0}{K_1 + K_2} \cdot 1 - E^{-(K_1+K_2)t} \tag{4}$$

Or when precipitation is complete

$$N = \frac{K_1}{K_1 + K_2} \cdot c_0 \tag{5}$$

or

$$\frac{N}{c_0} \quad \frac{K_1}{K_1 + K_2} \tag{6}$$

The fact that the rate at which the rate of snowfall decreases with time is not greatly influenced by temperature, while the total number of crystals produced increases by a large factor with a small temperature decrease, indicates that at higher temperatures K_2 is far greater than K_1. This is another way of saying that at higher temperatures the rate at which the silver iodide particles form snowflakes is so small that a large majority of them precipitate on the container walls before they have a chance to form snow crystals. At −20°C. the number of snow crystals obtained approaches the number of particles determined from the electron microscope, so that it is reasonable to assume that at this temperature most of the smoke particles form snowflakes.

If we assume that the rate at which precipitation decreases with time at the higher temperature is determined solely by K_2, then K_2 computed from the curve in figure 4 is approximately 6×10^{-3} per second. From the data in figure 5, N_1 and c_0 can be estimated and from them and the above value for K_2, values for K_1 for different smokes at different temperatures can be computed.

These values expressed as half-life in hours are as follows:

SILVER IODIDE	RANGE OF PARTICLE DIAMETER	HALF-LIFE	
		−13°C.	−10°C.
mg./second	*Å.*	*hours*	*hours*
1.5	30–700	5.5	173
7.5	50–1070	1.9	75
37.5	130–1400	1.4	67

If these data are extrapolated on the assumption that the log of the nucleation changes linearly with the reciprocal of the absolute temperature, we find that at −5°C. the half-life is several years, while at −20°C. it is a few seconds, and at −25°C. a few milliseconds.

This large effect of temperature on the nucleation rate is of the same order of magnitude as that observed in the measurements on supercooled water drops and on supercooled tin drops.

The curves in figure 5 show a maximum for the number of nuclei produced per gram of silver iodide at a certain rate of silver iodide consumption. This is probably because at a low rate of consumption, the particles produced are so small that their reaction rate is small and many of them precipitate out before they have a chance to serve as nuclei. At high rates of consumption, the particles formed are larger and react more rapidly, but because they are larger, fewer are formed per gram of silver iodide.

Diffusion coefficients of smoke

An average diffusion coefficient for the silver iodide smoke can be calculated from measurements made on the rate at which the smoke precipitates on the walls of a container. In the experiments which were made to find the rate at which the number of nuclei falls off as a function of the time the smoke is held in the sampling syringe, it was found that the half-life of the smoke was about 20 min. This was for a smoke produced when the generator was consuming 20 mg. of silver iodide per second.

Langmuir (2) has derived the following expression for the rate of precipitation of a smoke of diffusion coefficient D on the walls of a container having a volume V and an internal area A with slight convection caused by a slightly higher temperature at the bottom than the top:

$$\frac{d \ln c}{dt} = \frac{0.64 A D^{2/3}}{V} \tag{7}$$

where c is the concentration of the smoke at any time t.

Using this equation the diffusion coefficient is calculated to be approximately 6×10^{-5}.

Langmuir (2) has also given the following expression for the diffusion coefficient as a function of particle radius:

$$D = \frac{2.04 \times 10^{-16}}{a^2} + \frac{1.18 \times 10^{-11}}{a} \tag{8}$$

where a is the particle radius.

According to this expression, the diffusion coefficient 6×10^{-5} corresponds to a particle diameter of 400 Å. This is in reasonable agreement with the measurements from the electron microscope, which show this smoke to have a particle diameter ranging from 30 Å. to 1400 Å. with a median diameter of about 300 Å.

As has been shown, K_2, or the rate of disappearance in the cold chamber of silver iodide particles from causes other than nucleation, is about 5.7×10^{-3} per second.

Using the diffusion coefficient of 6×10^{-5} and equation 7, it can be calculated that the smoke concentration should fall at a rate of 1.3×10^{-4} per second. This is far less than the observed rate of decrease, K_2, which was found to be 5.7×10^{-3}. This calculation, of course, does not take into account the thermal diffusion caused by the vaporizer heater in the cold chamber, which would be expected to increase greatly the rate of precipitation on the walls.

Using the diffusion coefficient it is possible to estimate the rate at which the concentration of the smoke decreases because of precipitation on the water drops of the supercooled cloud. It can be shown that the rate of change of concentration due to this cause is given by the relation

$$\frac{\mathrm{d} \ln c}{\mathrm{d}t} = -4\pi D r_0 Q \tag{9}$$

where r_0 is the radius of the water drops and Q is the number of drops per cubic centimeter.

If we assume the liquid water contents to be 1 g. per cubic meter and the drop radius to be 5 microns, the rate of change of concentration from this cause is 5.6×10^{-4}.

SUMMARY

Measurements have been carried out on the nucleation of supercooled water clouds by silver iodide smokes. The smokes were produced by spraying a solution of silver iodide into a hydrogen flame. The particles of the smokes were found by electron microscope examination to range in diameter from 30 Å. to 1400 Å. The number of ice crystals produced per gram of silver iodide was determined as a function of the temperature of the supercooled cloud and the rate of introduction of silver iodide into the flame. Yields of ice nuclei of as high as 10^{16} per gram of silver iodide were obtained when the supercooled cloud was at −20°C. At −10°C. the same smoke produced only $10^{12.7}$ ice crystals per gram.

It has been found that silver iodide particles do not react immediately to form ice crystals when they are put into a supercooled cloud. Ice crystals were found to be still forming at a measurable rate 50 min. after a silver iodide smoke was introduced into a supercooled cloud. It is believed that fewer ice crystals are produced at higher temperatures than at lower temperatures, because the silver iodide particles react more slowly to form ice crystals and most of them precipitate on the walls of the cold chamber before they have time to react.

According to this interpretation of the results, the rate of reaction at −13°C. is thirty or forty times that at −10°C.

The author wishes to thank Dr. Irving Langmuir and Dr. Vincent J. Schaefer for their very helpful counsel and suggestions in this work. He is grateful to Mr. E. Fullam for making the electron microscope examinations and photographs and to Mr. Kiah Maynard and Mr. Duncan Blanchard for invaluable help in testing the smokes.

The research reported in this paper was made possible through support extended the Research Laboratory of the General Electric Company jointly by the Signal Corps and the Navy Department (Office of Naval Research) under Signal Corps Contract No. W-36-039-SC-32427.

REFERENCES

(1) Hill, A. E.: J. Am. Chem. Soc. **30**, 68 (1908).
(2) Langmuir, I.: Unpublished work.
(3) Schaefer, V. J.: Science **104**, 457–9 (1946).
(4) Schaefer, V. J.: Bull. Am. Meteorol. Soc. **29**, 175–82 (1948).
(5) Vonnegut, B.: J. Applied Phys. **18**, 593–5 (1947).

THE FORMATION OF ICE CRYSTALS IN THE LABORATORY AND THE ATMOSPHERE[1]

VINCENT J. SCHAEFER

General Electric Research Laboratory, Schenectady, New York

Received October 18, 1948

This paper summarizes some of the author's laboratory and field activities in the realm of experimental meteorology.

Some properties of natural snow particles in the free atmosphere and their electrical characteristics observed in snowstorms are given. A method for classifying snow particles is listed and the typical varieties illustrated by photomicrographs.

Methods of preparing plastic replicas of snow crystals and of snow on the ground are described. Techniques for detecting sublimation nuclei in the free atmosphere are given, involving the use of a cold chamber and supercooled films supported on rings or on thin plastic films.

A description is given of the technique used in simulating natural clouds in the laboratory, and the basic procedures which should be followed are listed.

A brief description of the rôle that certain types of dust particles play in serving as sublimation nuclei is mentioned and the results observed with a considerable number of typical samples and their effectiveness at various temperatures.

The effect on the crystal habit of ice produced by the blocking action of adsorbed surface-active polar molecules is shown.

The latter part of the paper describes briefly some of the aspects of the cloud modification studies initiated by the author in 1946 and subsequently carried out as Project CIRRUS, a joint Army–Navy–Air Force–General Electrical research project concerned with the physical nature of precipitation as snow and rain. A few of the observed results are mentioned, as well as the relationships which are believed to exist in unstable clouds.

During the past seven years, various phases of snow studies have been carried out by the writer and several of his associates of the General Electric Research Laboratory at Schenectady, New York. The earlier work was conducted in a snow laboratory on a slope of the Mohawk Valley in eastern New York. Subsequent studies starting in 1943 and continuing at the present time were also conducted at the Research Laboratory at Schenectady, New York, and the Mt. Washington Observatory in the state of New Hampshire.

The work has included such studies as the electrical properties of snow, the physical effects causing precipitation static on aircraft radio, various techniques for making snow crystal replicas, studies of the changes in snow on the ground, the physical processes related to the deposition of ice on aircraft surfaces, a study of supercooled water droplets, both in bulk phase and as an aerosol, methods for detecting sublimation nuclei in the free atmosphere, and finally, various phases of cloud physics, particularly "seeding" techniques which may be used to "trigger off" unstable cloud systems.

---

[1] Presented at the Symposium on Aerosols which was held under the joint auspices of the Division of Physical and Inorganic Chemistry and the Division of Colloid Chemistry at the 113th National Meeting of the American Chemical Society, Chicago, Illinois, April 22, 1948.

Also presented at the meeting of the International Commission of Snow and Glaciers as part of the Eighth General Assembly of the International Union of Geodesy and Geophysics, which was held at Oslo, Norway, August 27 and 28, 1948.

Neither space nor time permits more than brief descriptions of the various subjects which have been or are being studied at the present time. References are made, however, to papers which are available in the literature or as reports issued by our laboratory which may be used to obtain more detailed information on the various subjects discussed.

Fig. 1. Satisfactory exposure for point collector to record atmospheric electricity

THE ELECTRICAL PROPERTIES OF SNOW

Early in 1943 several of us in the Research Laboratory were asked to make a study of some of the causes of precipitation static. This is a trouble experienced when aircraft fly through various types of precipitation. Under such conditions, and particularly when the precipitation is in the form of snow, much static noise develops on the aircraft radio. Under severe conditions, the radio becomes useless because of the high noise level, and at times St. Elmo's fire becomes prevalent on the wing and propeller tips and other parts of the airplane.

To understand certain phases of the phenomenon better, the writer made a study of the electrical characteristics of snow particles and snowstorms. The general results of this research have been published (13).

In brief, the studies indicated that one of the main causes of static noise is the charging effect produced when snow particles hit a plane and are shattered, the fragments rubbing and bouncing across the metal surface thus serving as units of a frictional electrical generating system. The studies showed that a single snow crystal may be broken into thousands of fragments. As the plane becomes charged, it finally reaches an unstable condition at which time an electrical discharge occurs which may assume the form of St. Elmo's fire, as previously mentioned. These electrical discharges account for the radio noise.

TABLE 1

Comparison of general properties of positive current and cross current storms

GENERAL PROPERTIES	TYPE OF STORM	
	Positive current	Cross current
Air temperature	0–25°F.	25–35°F.
Precipitation forms	3, 4, 5, 7, 8	1, 2, 6, 8, 9
General nature of snowfall	Dry	Wettish
Clumping tendency	Rare	Common
Frozen water content	0.05–0.2 g./m.^{3}	0.05–1 g./m.^{3}
Rate of snowfall	0.1–1 cm /hr.	0.1–3 cm./hr.
Density on ground	0.05–0.15	0.05–0.75
Weight of crystals	4–100 micrograms	4–1000 micrograms
Falling velocity	50 cm. sec.$^{-1}$	20–250 cm. sec.$^{-1}$
Range of falling velocity	Small	Large
Rime deposition	Rare	Common
Cloud structure	Non-convective	Convective
Air-to-ground current	Positive	Positive and negative
Observed atmospheric electricity, C_p collector	Less than 0.25 μa.	-7.3 to $+9.7$ μa.
Frequency of electrical activity	Sporadic and short	Nearly continuous
Relationship of electrical activity to specific clouds	Not obvious	Often quite definite
Rain forms	Light	Thunderstorms

Among the interesting effects observed during the study of the electrical properties of snowstorms was that two types of snowstorms could be identified in terms of their electrical properties. These were termed cross current and positive current storms, the evidence indicating that the former were storms having clouds with a considerable amount of convective activity, while the latter storms apparently consist of relatively stable clouds.

The electrical activity during storms is easily measured by mounting a well-insulated collector such as a needle in a well-exposed position, as shown in figure 1. If the needle is connected by shielded cable to a recording microammeter and then to the ground, records such as shown in figure 2 are obtained in practically all snowstorms. After studying many storms, Table 1 was developed

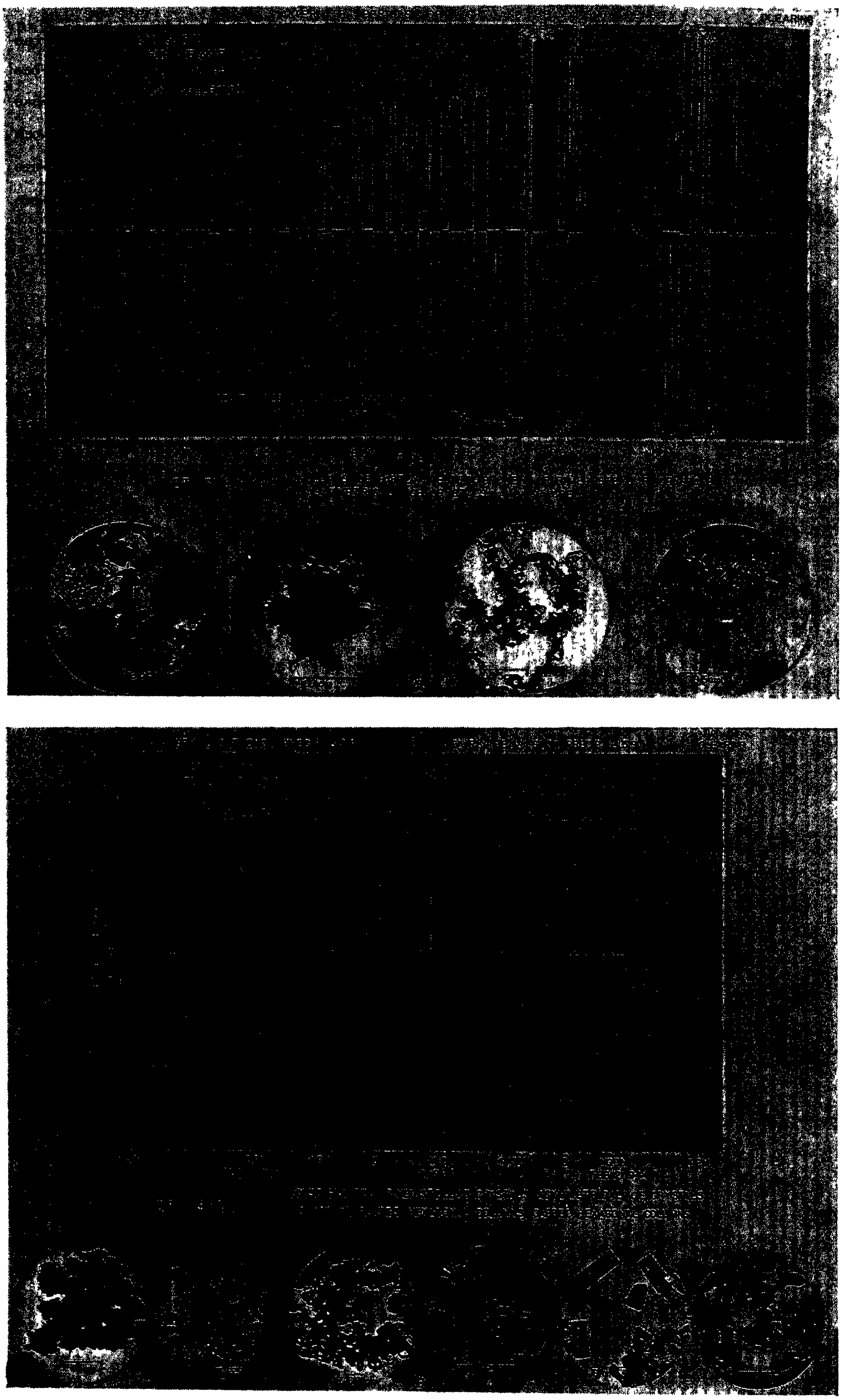


FIG. 2. Typical atmospheric electricity from positive current and cross current snowstorms.

to show the general characteristics of the storms producing these two interesting types of atmospheric electricity.

THE FORMATION OF ICE ON AIRCRAFT

One of the interesting developments following our study of precipitation static at the Mt. Washington Observatory was a shift of our interest from that subject to one of even more importance—the icing of aircraft.

Since Mt. Washington is noted for having "the worst weather in the world," it was found that the summit was an ideal spot to study some of the basic problems related to the icing of surfaces from supercooled clouds. Studies at the Observatory have been carried out each season since 1943 in an effort to understand better some of the fundamental phenomena. Much of the results of this work has been published in reports (11), most of which are available from either the U. S. Air Forces or our Laboratory.

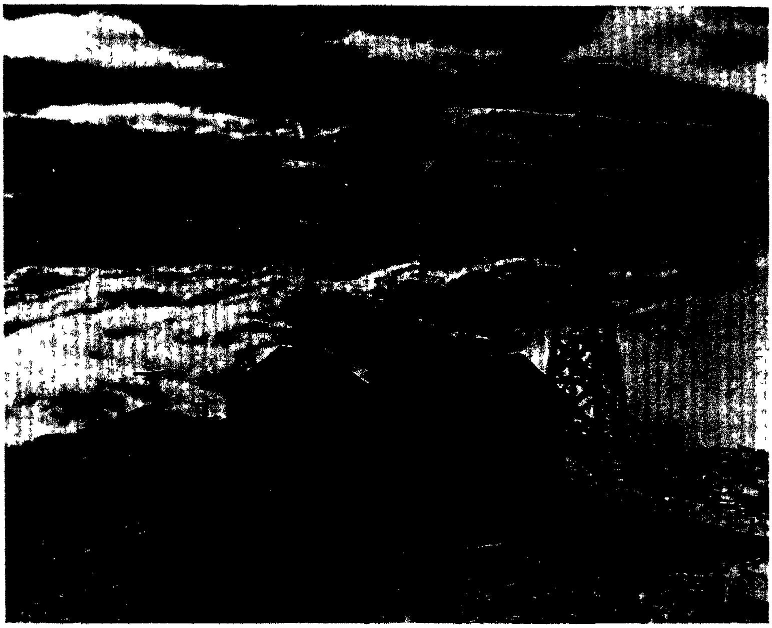


FIG. 3. View of Mt. Washington Observatory and supercooled clouds

It was the writer's visits to Mt. Washington and many observations made on the arduous climbs and descents of the mountain during icing storms and snowstorms that eventually led to the studies resulting in cloud modification techniques. A view of the Observatory after a mild icing storm is shown in figure 3. The clouds above and below the summit are all in a supercooled condition. Under somewhat similar cloud conditions but later in the season, figure 4 shows a relatively thin orographic cloud system which shed snow continuously for several days. A consideration of these observations and many others laid the groundwork for subsequent discoveries.

THE CLASSIFICATION OF SNOW PARTICLES

Any fundamental work with snowstorms and snow crystals immediately points out the necessity of having a simple classification chart to record the crystal forms occurring at a specific time from a storm.

Other investigators in this field (1, 2, 5, 8) have suggested classification methods, but a survey indicated that they were either more complicated than necessary or else failed to include one or more important types. The classification chart shown in figure 5 was prepared by the writer in 1943 for a group of co-operative observers who assisted him in making a study of the distribution of crystal forms which occur during widespread snowstorms in the northeastern part of the United States. This has been revised several times as the result of experience in the field, but in its present form has now been successfully used for more than three years.

FIG. 4. View of snow clouds from summit of Mt. Washington, New Hampshire

In order to describe the crystal forms observed at a particular time in a snowfall, the code number related to the crystal type is preceded by the letter P. If the crystal is coated with cloud particles so that it has a frosted appearance, the letter R following the code number denotes that condition. Thus the designation P 1 R, 6, 7 indicates that at the observation time stellar crystals coated with rime, ice needles, and asymmetrical crystals were falling from the sky.

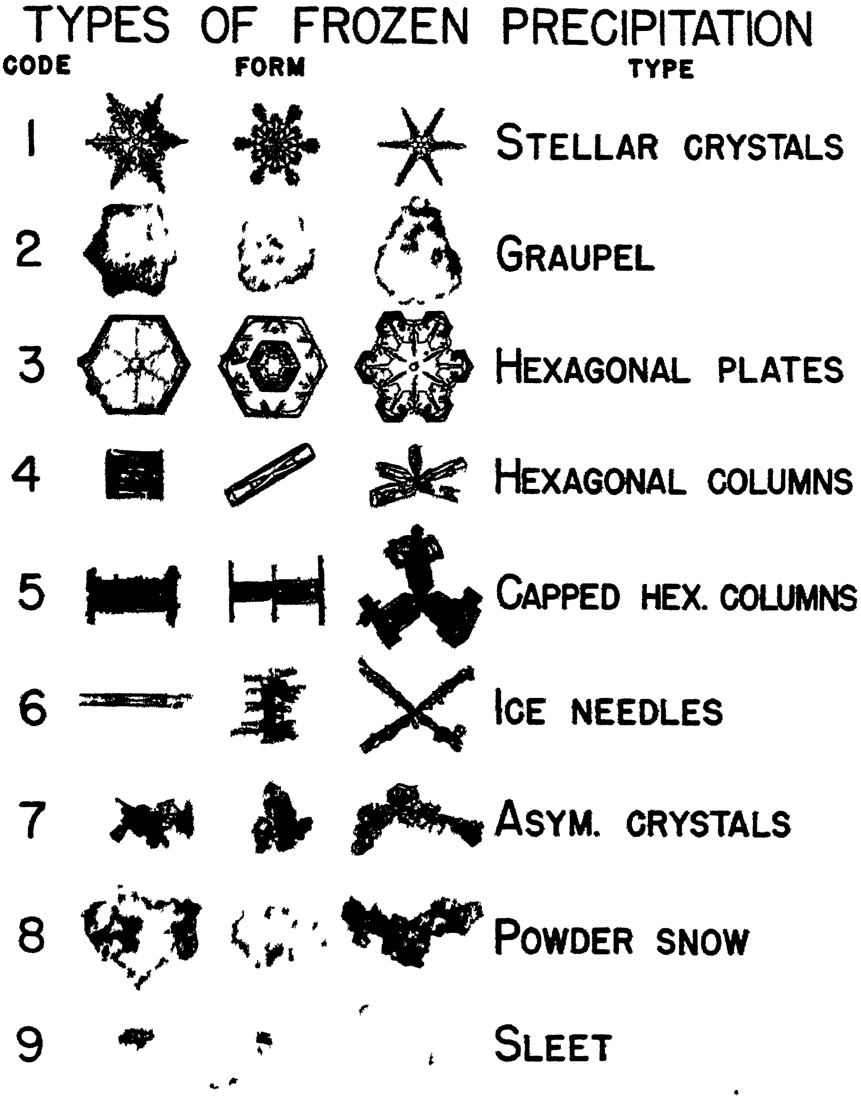


Fig 5 Classification chart of various forms of frozen precipitation

Greater detail may be included by affixing a subscript number indicating by tenths the relative quantities of each type at the observation time. The photomicrographs used to illustrate the chart show some of the variations which occur with each general type of snow particle. Table 2 indicates the general features and relative sizes of the nine different forms of precipitation. Figure 6 shows the range in size which is commonly observed with a typical crystal form—the hexagonal plate. The larger one comes from relatively low level clouds containing moderate quantities of moisture. Such clouds are generally below the 10,000 ft. level. The crystal shown in the middle of the figure is common to cirrus type clouds and may fall from an altitude as great as 40,000 ft. The smaller crystal is the type known as diamond dust, which commonly forms when the air is supersaturated with respect to ice. Such conditions generally happen

TABLE 2

General features and relative sizes of frozen precipitation

1. Stellar crystals.	Six-rayed starlike forms occurring as single crystals or in clumps; crystals vary in a wide range of form and size, often comprising cottony flakes sometimes 2 in. in diameter; size $\frac{1}{32}$–$\frac{1}{2}$ in.; clumps common
2. Graupel.	Snow crystals which are covered with frozen cloud particles forming a rime deposit; this results in pellets of soft snow, which range from hexagonal to rounded kneaded forms; size $\frac{1}{16}$–$\frac{1}{4}$ in.; rarely clumped
3. Hexagonal plates	Thin, hexagonal, solid, or semisolid plates often containing internal structure due to air inclusions; the smallest form of this crystal and the stellar form is known as "diamond dust"; size $\frac{1}{1000}$–$\frac{3}{16}$ in.; rarely clumped
4. Hexagonal columns	Generally flat-ended, sometimes with one end pointed, hexagonal prisms mostly of transparent ice, many with air inclusions; often a group of columns grow from a common source radiating in several directions; the former type produces 22° and 46° halos around the sun; size $\frac{1}{64}$–$\frac{1}{8}$ in.; sometimes clumped
5. Capped hexagonal columns ..	Similar to hexagonal columns with the exception that both ends terminate as expanded hexagonal plates; a hexagonal plate sometimes is positioned midway between the two plates; many queer forms occur; size $\frac{1}{64}$–$\frac{3}{16}$ in.; sometimes clumped
6. Ice needles	Long slender shafts with hexagonal cross-section often terminating with sharp irregular points on periphery; needles occasionally grow in masses in random directions; size $\frac{1}{64}$–$\frac{3}{8}$ in.; often clumped
7. Asymmetrical crystals	Angular crystals without a symmetrical outline occurring as simple or compound crystals; sometimes crystalline rays originate from a common center but grow in many directions; size $\frac{1}{64}$–$\frac{1}{8}$ in.; often clumped
8. Powder snow.	Dry bits of snow without angular form, often of irregular shape; size $\frac{1}{64}$–$\frac{1}{8}$ in.; rarely clumped
9. Sleet	Pellets of translucent to transparent ice often bearing bumps or other protuberances; they are frozen rain drops but are not always round; size $\frac{1}{32}$–$\frac{1}{4}$ in.; very rarely clumped

with strong temperature inversions accompanied by cloudless skies, the crystals coming from a region less than 1000 ft. above the ground.

COLD CHAMBER STUDIES

The basic experiment (12, 16) in which a supercooled cloud of liquid water droplets may be converted to a mass of snow crystals is a spectacular one which may be easily reproduced with very simple apparatus.

The essential articles for the experiment are an open-topped, well-insulated chamber having a volume of about 1/20 cu.m., e.g., 48 cm. long, 24 cm. wide, 48 cm. deep, which may be cooled to a temperature of −10° to −20°C. The walls and floor of the chamber should be painted dull black or lined with black velvet, cloth, or cardboard.

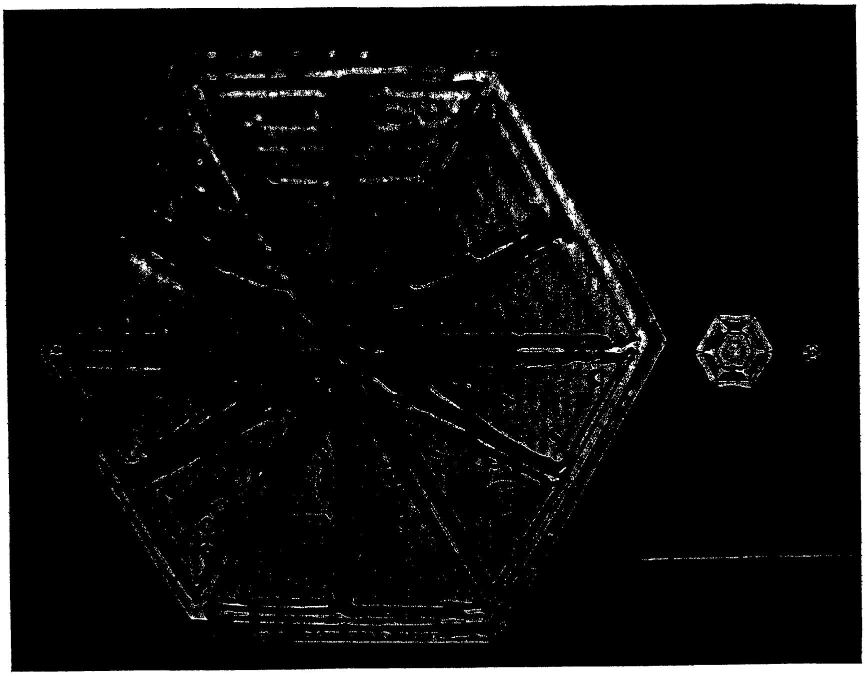


FIG. 6. Relative sizes of hexagonal plate type of snow crystals

Illumination is simple but important. For best results, the light should produce a brilliant beam and is most satisfactory when mounted above the chamber and directed into it at an angle of 45° from the horizontal.

A supercooled cloud is produced by placing a moist cloth in the cold air of the chamber. The moisture condenses into small water droplets, generally in the size range of 10–20 microns diameter. Another simple method for forming the cloud is merely to breathe into the cold air. Several breaths will produce a cloud having a liquid water content equivalent to 1 g./cu.m. A chamber with a cloud in it is shown in figure 7.

Under ordinary conditions, few crystals will be seen in the chamber. A single

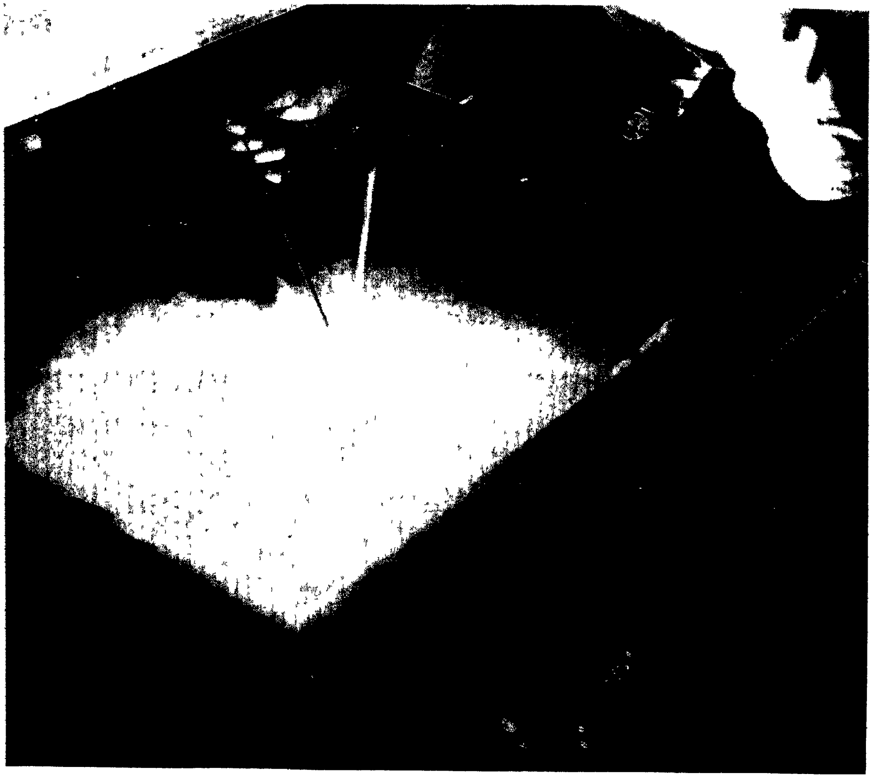


FIG. 7. View of cold chamber containing supercooled cloud

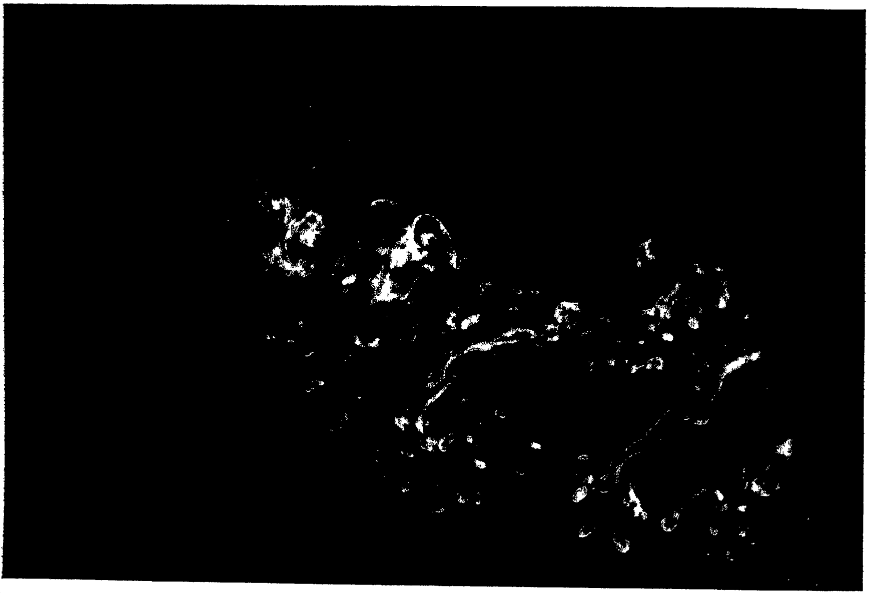


FIG. 8. Trail of ice crystals produced in cold chamber by seeding with needle cooled below −39°C.

crystal may easily be seen among the water droplets, since it grows rapidly, scintillating in the light. They are most easily seen when viewed at right angles to the axis of the beam, since much of the light from them is specular reflection. This is particularly true of hexagonal plates, since they become oriented with their flat surfaces horizontal.

It can easily be shown that most of the condensed cloud formed in the chamber is supercooled. If a fine wire is rotated at 180 R.P.M. in the chamber, it becomes coated with ice. If left alone, the cloud slowly disappears by condensing onto the cold walls of the chamber in the form of frost.

FIG. 9. Twinkling crystals

Within a minute after the supercooled cloud has been formed, the crystals which develop from sublimation nuclei present in the air grow and fall to the floor of the chamber. The supercooled cloud may be suddenly converted to snow crystals by briefly introducing an object colder than −39°C. This may be a needle cooled in liquid air or a tiny particle of dry ice. A single passage of such a cold object produces an effect (figure 8) which quickly changes to a mass of twinkling crystals like those in figure 9. These crystals range in size from 10 to 200 microns. By the time they reach this latter size, the falling velocity is of the order of 1 cm. sec.$^{-1}$, so that it is difficult to keep the particles from "snowing out." Typical crystals produced in this manner are shown in figure 10.

One of the most convenient ways to convert supercooled clouds to ice crystals is by the use of dry ice. Dry ice is easily produced by permitting liquid carbon dioxide to expand into the air. A very convenient source of dry ice for experimental studies is the small 10-g. shells, such as used in soda fountains and in various home appliances. A simple valve for controlling the opening of such containers is available as part of a small model-airplane carbon dioxide engine. By valving the container held in an inverted position above the cold chamber, tiny bits of dry ice will be seen to fall through the cloud, leaving behind tiny white "contrails." These rapidly grow until the cloud is completely replaced by a mass of twinkling ice crystals.

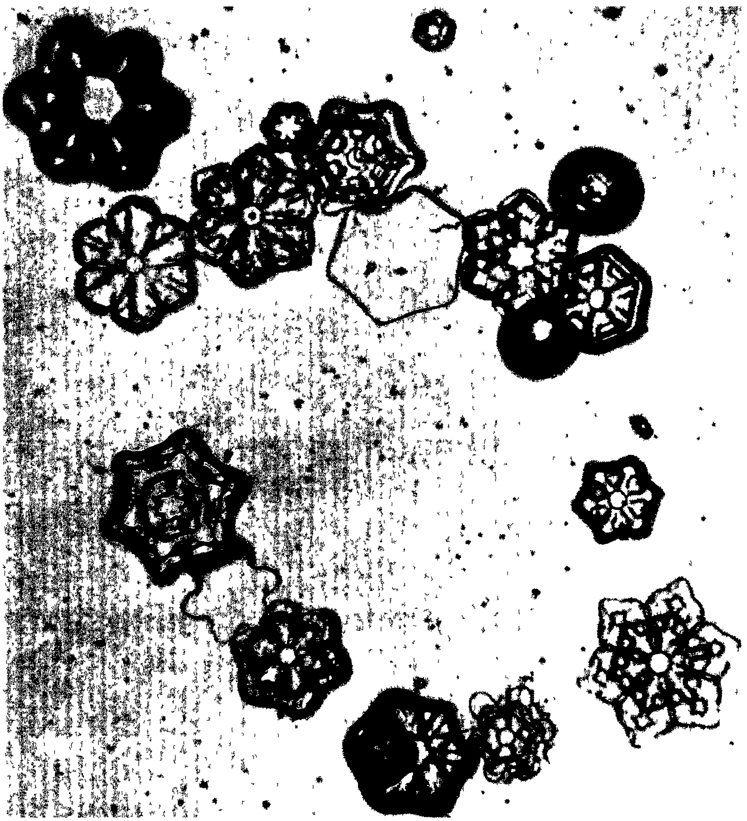


Fig. 10. Typical photomicrograph of snow crystals produced in cold chamber

Several very interesting phenomena may be seen in this experiment. The cold chamber, when operated with the lid removed, has a temperature gradient with the air warmest at top, colder below. This temperature inversion produces a very stable air condition in the chamber, which resists convection and turbulence. Thus when seeded with dry ice or a cold needle, the heat released by the crystallization process is not sufficient to overcome the inversion stability. If the region of ice crystals is observed carefully, a dark area quickly develops at the interface between them and the supercooled cloud. This is caused by the fact that the ice crystals grow at the expense of the water droplets which evaporate, the ice crystals growing from the condensation of gaseous water vapor, since the vapor pressure of ice is below that of water at all temperatures below 0°C. The dark area often grows to a width of 1.5–2 cm. in a few seconds, showing in a striking manner the phenomenon of diffusion. Another interesting and unusual effect is observed whenever the concentration of ice crystals in the chamber exceeds about 500 per cubic centimeter. In addition to the glint from the crystals as they tumble about in the chamber, a background twinkling is observable. This is due to Brownian motion and is an elegant way to see this interesting phenomenon.

Many other experiments could be described which may be carried out in the cold chamber. These must be deferred at this time, but it should be emphasized that the cold chamber as described is so simple to use that it should become a familiar research tool and demonstration unit for those working in meteorology and other related physical studies.

THE PREPARATION OF SNOW CRYSTAL REPLICAS

Excellent replicas of snow crystals may be prepared by a very simple method which has been described elsewhere (9, 10).

Briefly, it consists in using a dilute solution of a resin in a solvent which does not dissolve ice. An excellent solution consists of 1 per cent polyvinylformal dissolved in ethylene dichloride. When cooled below 0°C., it may be used to produce exact replicas of snow crystals by letting natural crystals fall on a collecting sheet of black velvet cloth or paper and then placing them in a small drop of solution previously placed on a piece of cold glass or slightly absorbent black cardboard. A variation of the technique is to coat a cold glass plate or black cardboard sheet with a thin film of the replica solution, which is then exposed to the falling snow crystals for a specific time period ranging from 5 to 60 sec., depending on the density of the snowfall. This produces a permanent record of the snowfall at that particular time in terms of snow crystal type, concentration of crystals, and other factors of importance to obtaining a better understanding of the physical features of such storms.

The technique is so easy that relatively untrained observers can obtain excellent results by observing a few basic rules. The method also lends itself admirably to sampling from planes, at remote stations, and by automatic exposure instruments. The samples thus obtained may then be studied at leisure in a warm laboratory. A photomicrograph of a replica is shown in figure 11.

THE PREPARATION OF REPLICAS OF SNOW ON THE GROUND

In addition to the various replica techniques previously described, there is another which may be used to advantage in preparing replicas of snow on the ground. For this purpose, a more concentrated solution of polyvinylformal should be used, having a concentration of between 3 and 5 per cent.

To illustrate one of the possibilities, replicas were prepared of a cross section of snow on the ground, as shown in figure 12. To prepare these records, a 5-cm. strip of 3 per cent solution was spread lengthwise on a piece of window glass 24 in. long. A vertical face was cut into a fresh blanket of snow and the coated glass pressed firmly against it, withdrawn, and then placed in a well-ventilated but sheltered location to permit the solvent to evaporate. Until all traces of the solvent have evaporated, the replica must be kept at a temperature colder than 0°C. The best replicas are obtained when the snow particles are permitted to sublime. If it is desirable to hasten the process, however, the sample may be warmed as soon as the solvent is gone, which may take from 20 to 60 min. depending on the ventilation and the temperature of the replica sample. After obtaining a cross-section replica as described, subsequent samples may be obtained by successive applications of the solution side by side until the surface area of the backing is exhausted.

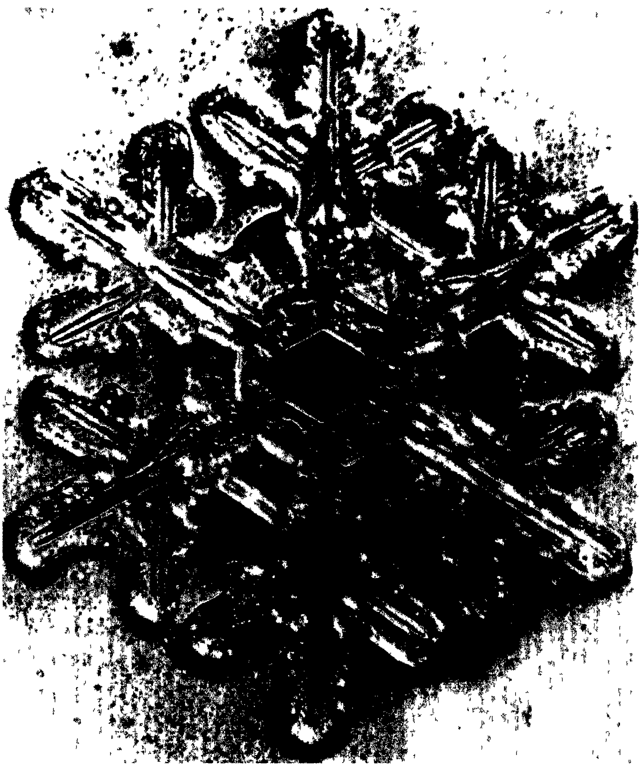


FIG. 11. Photomicrographs made by transmitted light of snow crystal replicas

The sample shown in the photograph illustrates a vertical shrinkage of the snow layer during a period of 5 successive days during which time the air temperature remained continuously below 0°C. Although the shrinkage amounted to more than 50 per cent in snow thickness, the density increased from less than 0.10 to more than 0.20, so that the content of liquid water remained nearly the same. Examination of the structure of the snow particles on the fifth day showed that, although most of the original snow consisted of perfect crystals of stellar, hexagonal plate, and column types, a major change occurred during the 5-day period, producing granular snow approaching the rounded firn forms.

Subsequently, as shown by the photograph, several additional light snowfalls occurred which were also recorded by the replica technique.

It is hoped that eventually a snow replica solution may be developed employing the solventless type of plastic in which several monomers are catalyzed by an oxidizing agent to form a true three-dimensional structure. Until such a material becomes available which employs materials which do not modify ice, satisfactory replicas may be prepared as described, particularly when good ventilation is provided.

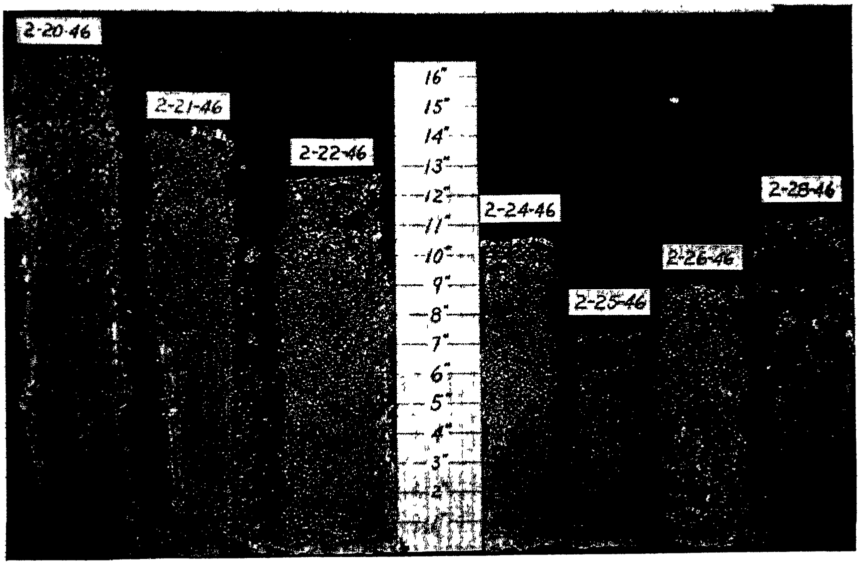


FIG. 12. Replica of cross section of snow deposit on ground

THE DETECTION OF ICE NUCLEI IN THE FREE AIR

The random occurrence of supercooled clouds in the free atmosphere is a well-known fact. The reason for this seemingly haphazard phenomenon is probably closely related to the concentration of sublimation nuclei in the particular air mass when clouds occur.

A better knowledge of sublimation nuclei is needed, particularly of particles other than ice which lead to the formation of ice crystals in air supersaturated with respect to ice. Unfortunately, such particles are generally of microscopic or submicroscopic size and, in addition, include only a small fraction of the dust in a given sample of air.

It is important to differentiate between *condensation* and *sublimation* nuclei. The former are always much more common than the latter. As will be shown in a subsequent section, a third type of particle, the *freezing* nucleus, is also of importance in this connection.

Three methods have been devised by the writer for detecting sublimation nuclei in the clear air of the atmosphere (17).

The most satisfactory of these methods is to use the cold chamber of the type previously described. A sample of the surrounding air is introduced into the chamber by simply removing the air within it by employing a suction fan. Unless the air which is introduced has a very low humidity or was nearly as cold as that in the chamber, condensation will occur. By illuminating this newly formed cloud with a strong beam of light, every active sublimation nucleus which was present will appear as a twinkling snow crystal, easily discernible among the surrounding supercooled cloud droplets. The temperature of the air in the cold chamber will have a direct relationship to the number of particles likely to be observed.

Observations made during 1947 and 1948 on the summit of Mt. Washington (6288 ft.), New Hampshire, and in the Mohawk Valley show the concentration of sublimation nuclei at a cold chamber temperature of −20°C. to vary from none observable to values up to 10 per cubic centimeter. The type of crystal which forms under such conditions probably bears a definite relationship to the substance serving as the active nucleus. These crystals may be examined by direct examination under a microscope kept in the chamber, by preparing plastic replicas, or by permitting the crystals to seed a thin supercooled water film.

Certain "soap" bubble solutions serve admirably for this latter purpose. Most solutions work except those containing large amounts of glycerol or similar freezing-point depressants. A film is formed on a ring 5–10 cm. in diameter and then held in the cold air of the chamber. The film supercools in about a second. If visible crystals are in the air, the number of separate crystals which develop and grow in the film appears to be in direct proportion to that concentration. At least four types of crystals grow in the film. A six-membered stellar form appears when the crystals are hexagonal plates. A symmetrical cross or x-shaped crystal appears when hexagonal columns occur. An irregular-shaped crystal predominates when asymmetric crystals are present. It should be remembered that there may also be crystals developing from freezing nuclei within the solution. Whenever this occurs, the crystals tend to have frostlike forms with a tendency to be curved. There is a time lag also before such crystals appear, as compared to the effects seen when sublimation nuclei or ice crystals are present.

A variation of this method is to blow bubbles in the free air when its temperature is below freezing. When this is done, crystals will sometimes develop in the bubble. Occasionally a single crystal will grow completely around a bubble 10 in. in diameter. Figure 13 illustrates an effect commonly seen. Tables 3, 4, and 5 show typical data indicating the variations which have been observed during the past winter.

Permanent replicas of such effects may be prepared in the following manner: On a frame of convenient size, such as 9 x 12 cm., a thin film of cellulose acetate is stretched and cemented. This film is then coated with a thin layer of 3–5 per cent of polyvinyl alcohol dissolved in water. This solution readily supercools and when seeded with sublimation nuclei or ice crystals will produce effects

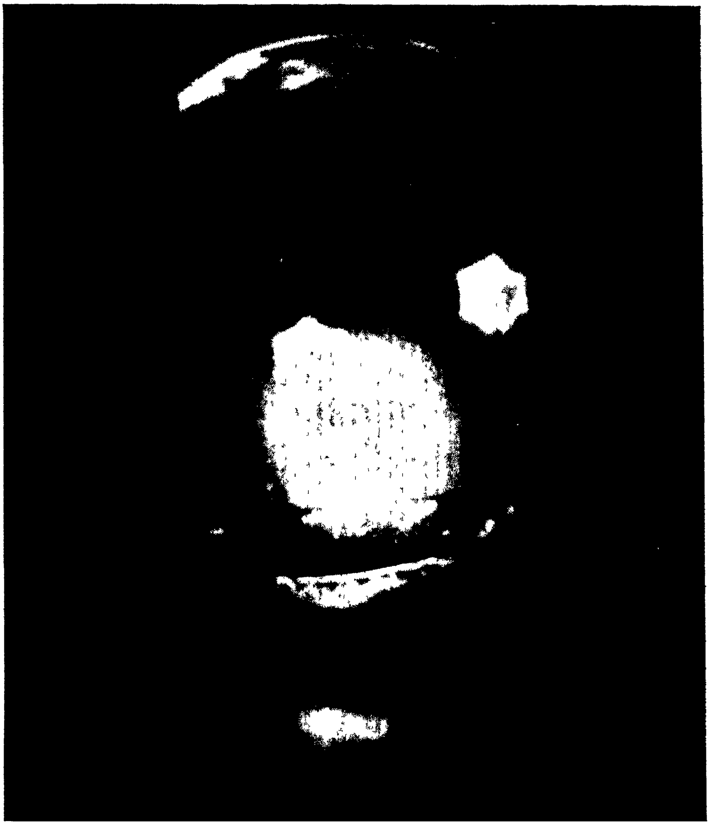


FIG. 13. Supercooled bubble in which ice crystals are growing

TABLE 3

Extremes in number of sublimation nuclei as a function of air temperature

DATE	TIME	TEMPERATURE	NUMBER OF NUCLEI	TYPE	COMMENTS
		°C.			
February 11, 1948	0800	−18	>50	Hexagonal	Alto stratus clouds
February 24, 1948	0700	−18	20–30	x	Supercooled fog in lowlands
December 25, 1948	0730	−14	30–40	Hexagonal	Clear
January 23, 1948	0800	−14	10–15	Hexagonal	Stratus overcast
January 22, 1948	0815	−10	10–20	Hexagonal, x	Stratus overcast
February 13, 1948	0745	−10	3–5	Hexagonal, x	Cirrus haze
January 25, 1948	1200	−7.8	20–30	Hexagonal, x	"Frosty" air
February 19, 1948	0730	−7.8	1–5	x	Cirrus haze
February 15, 1948	0800	−6.6	5–10	Hexagonal, x	0.8 strato cumulus
March 1, 1948	2200	−6.6	None		Cirrus haze
February 24, 1948	2300	−6.1	None		Clear, 22° halo
March 15, 1948	0645	−6.1	None		Cirrus haze

identical with those already described for the soap film. Figure 14 shows two samples of the growth of ice crystals in such films. If, after the crystallization is completed, the film is kept below freezing until the ice disappears by sublimation, a permanent replica remains. This was formed by the polyvinyl alcohol being precipitated at the interfaces between the ice crystals.

TABLE 4

Variation in number of ice nuclei under nearly isothermal conditions

DATE	TIME	TEMPERATURE	NUMBER OF NUCLEI	TYPE	COMMENTS
		°C.			
February 11, 1948	1300	−9.5	20–30	Hexagonal, x	Alto cumulus
January 22, 1948	0815	−10	10–20	Hexagonal, x	Thin cumulus humilus
January 23, 1948	0745	−9	10–20	Hexagonal	Cumulus humilus
February 23, 1948	0800	−9.5	5–10	Hexagonal, x	Scattered fracto cumulus, hazy
February 13, 1948	0745	−10	3–5	Hexagonal, x	Cirrus haze
March 11, 1948	0645	−9.5	2–5	Hexagonal	0.8 strato cumulus sun pillar
February 22, 1948	2200	−9.5	1–3	Hexagonal, x	Clear
March 6, 1948	2130	−9	1–2	Hexagonal, asymmetrical	Cirrus haze
March 13, 1948	2200	−9	1–2	x	Thin haze
February 20, 1948	0745	−9.5	0.1–1.0	Hexagonal	Clear

TABLE 5

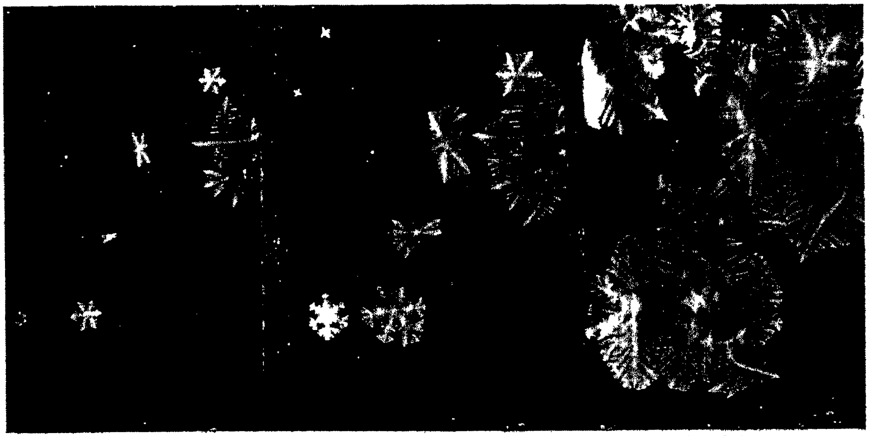
The relationship of concentration of sublimation nuclei to the temperature of the free atmosphere

Cold chamber method—Mt. Washington Observatory

TEMPERATURE	NUMBER OF OBSERVATIONS SHOWING VARIOUS CONCENTRATIONS			
	ca. 1 × 10^6/m.^{3}	ca. 1 × 10^3/m.^{3}	< 1 × 10^2/m.^{3}	None
°C.				
−30	4	5	1	1
−26	3	3	1	4
−22	5	4	0	1
−18	6	4	0	1
−14	3	5	0	3
−10	5	4	0	3
−6	1	2	5	4
−2	0	1	3	7
+2	0	3	3	6
+6	2	0	6	3
+10	0	2	3	6
+14	0	1	9	1

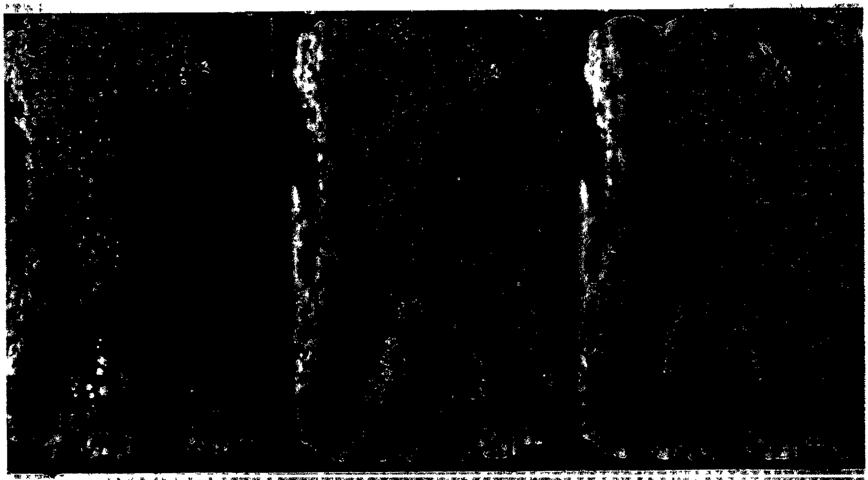
THE EFFECTIVENESS OF CERTAIN DUSTS AS SUBLIMATION NUCLEI

It is obvious that there are physical processes in the natural atmosphere which lead to the formation of snow at temperatures only a few degrees below 0°C. In the past, it has been suggested (3) that graphite and quartz dust carried into the atmosphere were the active sublimation nuclei for ice crystals. The hexagonal structure of these materials in their crystalline form was thought to be of importance in their activity. So far as I am aware, no information has been published of any verification of these ideas by laboratory experiments or observations of nuclei of actual crystals.



10 SEC. 20 SEC. 30 SEC.

EFFECT PRODUCED BY 10 NUCLEI



10 SEC. 20 SEC. 30 SEC.

EFFECT PRODUCED BY SEEDING WITH MANY NUCLEI

FIG. 14. Stages in growth of ice crystals in supercooled 3 per cent solution of polyvinyl alcohol.

It was found (19) by Dr. Vonnegut of our laboratory that small particles of silver iodide were quite effective as ice nuclei. This material has a rather high temperature coefficient, the warmest temperature at which any particles are active being about −5°C. with a much greater activity at −10°C.

During the past year, I have found a number of naturally occurring dusts which serve as effective sublimation nuclei. Among these are volcanic dusts, silicates, and clays. In most instances, these materials have been selected from extensive natural deposits in arid or semiarid regions noted for their duststorm activities. Each material has its own temperature range of activity for producing effective nuclei. For example, a sample of loess from an ancient lake bed near Richland, Washington, when formed as an airborne dust shows the presence of a small percentage of particles that form snow crystals in saturated air at a temperature of −12°C. At −24°C., all of the particles are "active" nuclei. Another dustlike soil from a desert in Nevada does not show any activity until the temperature reaches −24°C. and does not reach 100 per cent activity until the temperature is −33°C. The temperature–activity relationships of a number of such natural dusts are shown in figure 15 and several typical snow crystals grown on such foreign nuclei are shown in figure 16.

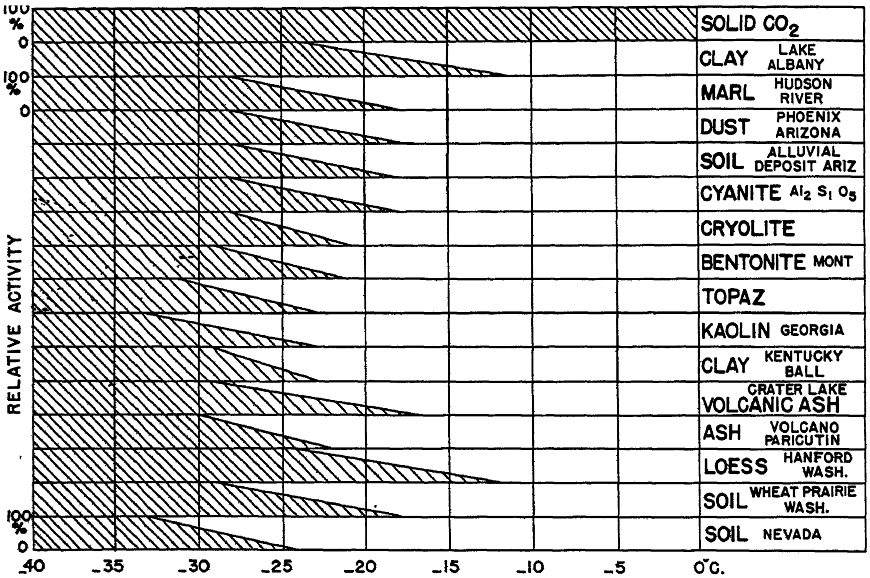


FIG. 15. The temperature–activity relationships of various kinds of foreign particle ice nuclei.

It is believed that the mechanism responsible for the activity of such particles is the fact that the surface of the particle is conducive to the formation of frost. As soon as a frost deposit occurs on the surface of the particle, growth continues as though the particle were made up only of ice. Dust particles larger than 1 micron generally produce asymmetrical crystals with irregular platelike growths, although smaller particles often lead to the development of very symmetrical ice crystals.

Since these dusts are active in the temperature range where snow is commonly observed, it is quite likely that they are responsible for the initiation of many snowstorms. It is unlikely, however, that such materials would be so concentrated that they could provide the tremendous quantities of nuclei necessary to keep a snowstorm going. It has been my observation that the majority of snow particles from storms coming from relatively low clouds are asymmetrical crystals which seem to be growing on fragments of other crystals. This strongly suggests that snowstorms might often be initiated by foreign particle sublimation nuclei in the early stage of development of the storm. These particles grow big enough and have a sufficient concentration so that for a short time they collide with one another. The resulting fracture from collision and abrasion and the electrical effects which result would then produce large numbers of tiny ice fragments, each of which would serve as an effective nucleus up to the freezing temperature. This mechanism is thus a chain reaction of a type somewhat comparable to that described by Langmuir (6) for the development of precipitation in warm clouds.

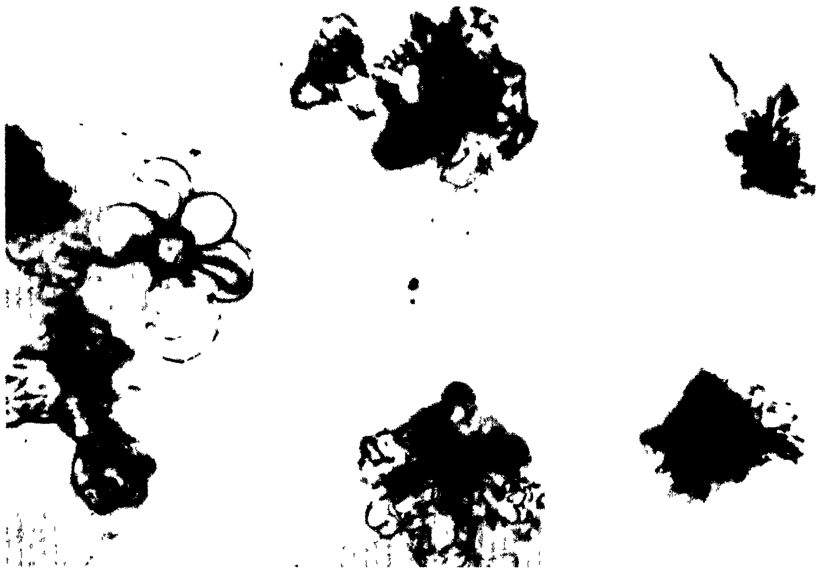


FIG. 16. Typical snow crystals formed on foreign particle nuclei

THE MODIFICATION OF SNOW CRYSTAL FORMS BY VAPORS

The variations in snow crystal forms shown in figure 5 are generally believed to be due to variations in temperature and moisture supply. It has been shown by Nakaya (7) that most of these variations in crystal form may be duplicated in a laboratory cold chamber by carefully controlling the temperature and moisture supply as a crystal grows on a supporting fiber.

Work in our laboratory (20) shows that modifications to the crystal form may also be produced by introducing certain organic vapors into the chamber in which the crystals grow. Unlike Nakaya's experimental method in which single crystals are grown, our experiments generally produce a minimum of 10^7 crystals.

A typical experiment employs air having a mean temperature of −15°C. Moisture is added to the air by the evaporation of water from a porous porcelain container or by use of a nebulizer (a small aspirator constructed within an outer glass container which directs an atomized water stream against the surface of the outer wall, the finer particles and accompanying saturated air emerging as a stream of water condensate from a tube on the top of the nebulizer).

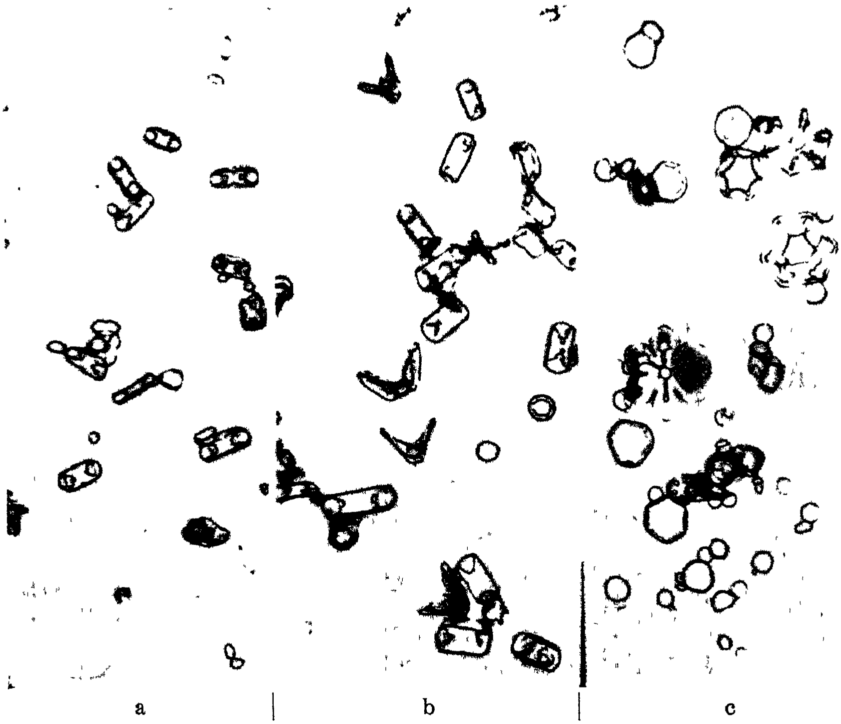


FIG. 17. The effect of acetone vapor on the crystal habit of ice. (a) Effect of 1 molecule of acetone with 10 molecules of water; (b) effect of 1 molecule of acetone with 100 molecules of water; (c) effect of 1 molecule of acetone with 1000 molecules of water.

Under normal conditions when the supercooled cloud is "seeded" with dry ice or when adiabatic expansion cools a small sample of the air below −39°C., the ice crystals which form in ordinary laboratory air are simple hexagonal plates. When, however, a small amount of certain organic chemicals is introduced into the air, either by vaporization or by use of the nebulizer, the crystals which form are modified. For example, as shown in figure 17, acetone in concentrations as low as one molecule of acetone to 100 molecules of water causes all of the crystals to become either hexagonal columns or a peculiar combination of plate and column. On the other hand, the introduction of the vapor of nitric acid or nitrous oxide (figure 18) tends to emphasize the trigonal symmetry of the water molecule so that the majority of the crystals are triangular with some of them being almost perfect isosceles triangles with sharp corners bearing no suggestion of the hexagonal pattern so common to crystals formed in "normal" air.

The use of acetic acid (figure 19) encourages the development of large stellar-like crystals containing intricate structural designs. Larger molecules of the fatty acid series having an appreciable vapor pressure, such as valeric and caprylic acids, produce the columnar form of prism similar to that observed with acetone.

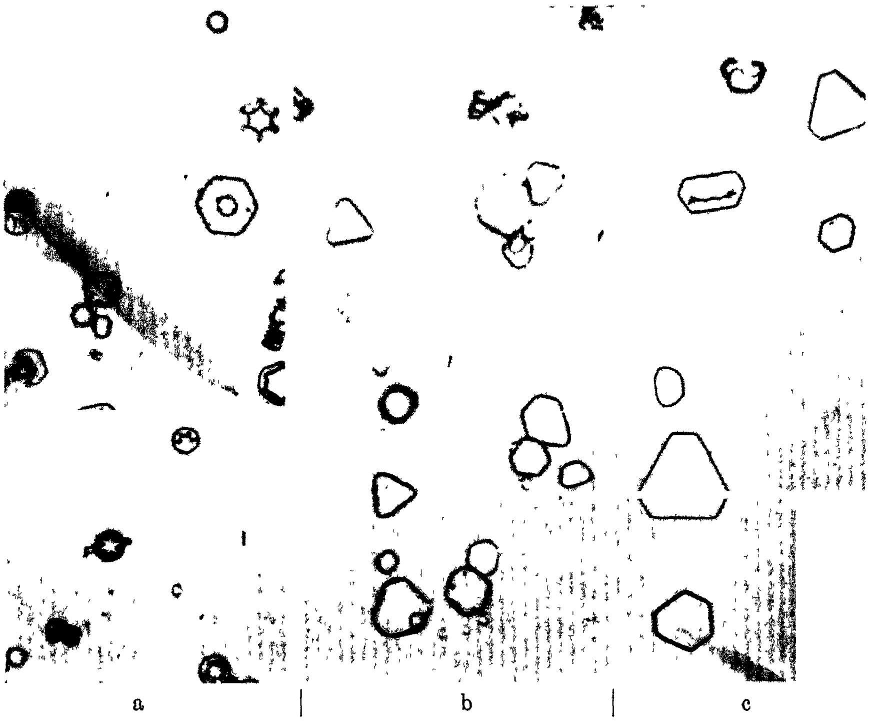


FIG. 18. The effect of nitrous oxide and nitric acid vapor on the crystal habit of ice. (a) Crystals formed in ordinary air; (b) effect of trace of nitric acid vapor in air; (c) effect of nitrous oxide smoke in air.

The aldehydes also modify the ice crystal structure, formaldehyde causing rather short columns while heptaldehyde produces a peculiar crystal somewhat akin to the capped hexagonal column.

Other chemical types which produce definite modifications are dioctyl sodium sulfosuccinate (OT 100), heptyl and octyl alcohols (methyl and ethyl are not effective), certain of the silicones, notably dimethyldichlorosilane, pentamethylaminomethyldisiloxane, and the ketones.

A preliminary consideration of the results shows that the most effective chemicals for producing modification are those which have strong polar properties. These are generally surface-active substances and presumably produce the modification by becoming preferentially adsorbed on certain faces of the crystals. This restricts growth in some directions, allowing more growth on others.

The study up to the present time indicates that the modification of the crystal is a continuing phenomenon and not one which is directly controlled by the crystal habit of the nucleus. By producing a batch of crystals by the adiabatic expansion of air containing acetone vapor and then permitting them to grow in saturated air under normal conditions, a typical crop of hexagonal plates was produced.

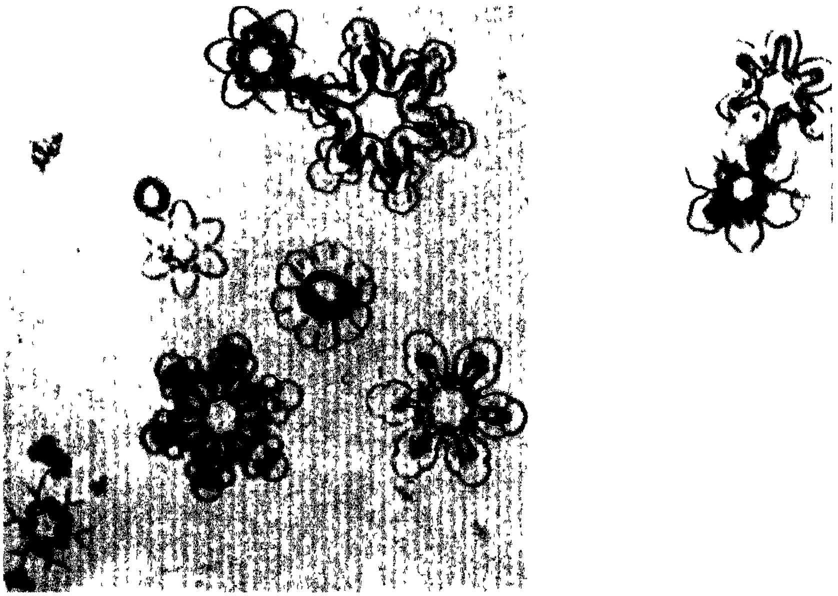


FIG. 19. The effect of acetic acid vapor on the crystal habit of ice

This experiment was made by placing a tiny wad of filter paper soaked in acetone into a "pop" chamber having a capacity of 0.025 cc. Saturated air at 25°C. was introduced into this chamber under a pressure of 30 lb. per square inch. By the sudden opening of the chamber, using a high-speed release system, the air expanded and cooled below the critical temperature of -39°C., causing the spontaneous formation of ice crystals in great numbers. These grew at the expense of a supercooled cloud in the cold chamber. Instead of forming hexagonal columns, as would be the case if the original crystal habit were dominant, the crystals which formed were of the "normal" variety of hexagonal platelets.

Work in this field is continuing, since many fundamental phenomena in the field of crystallography may be studied with great ease with the techniques which have been developed.

NOMENCLATURE IN EXPERIMENTAL METEOROLOGY

As the new science of experimental meteorology develops, especially as applied to cloud seeding and modification, certain experimental procedures are necessarily so often described that it becomes worthwhile to introduce new terms to simplify description and reduce the chance of ambiguity.

Since clouds occur in air which may be warmer or colder than 0°C., it was decided by our group to follow a suggestion of Langmuir's and adopt the following terminology.

The term a *cold cloud* shall be used to describe a cloud of water droplets which is supercooled and all of the cloud is in air which is colder than 0°C.

A *warm cloud* consists of a cloud of water droplets existing in air of which all parts are warmer than 0°C.

A *cool cloud* shall be the term used to describe a cloud which combines the properties of both *warm* and *cold* clouds; that is, the cloud exists in both above and below freezing air, with the freezing level extending through some part of it.

Other terms are being devised to describe some of the observable features peculiar to cloud modification studies. These will be described in greater detail in a later paper.

THE PRODUCTION OF SNOW IN SUPERCOOLED CLOUDS

As a result of the experimental studies made in the cold chamber during the summer of 1946 (12, 16), which has already been mentioned, the first experiment in the free atmosphere carried out by the writer took place on November 13, 1946. On that day a 4-mile supercooled alto stratus cloud having a temperature of -18.5°C. and containing no trace of ice crystals was "seeded" with 6 lb. of dry ice by flying through it in a small Fairchild monoplane at an altitude of 14,000 ft. Half of the fragments were put out from a mechanical dispenser operating on the floor of the plane; the rest were scattered by merely letting the slip stream suck the fragments from cardboard containers held out of the open window. Figure 20 shows the cloud before and after seeding.

Since that time many experiments have been carried out by the writer, by members of Project Cirrus—a joint Army–Navy–Air Forces–General Electric weather research program—, and by many other investigators. Some of these results may be studied by referring to detailed reports and papers (4, 14, 15, 18). Positive results are always observed when supercooled stratiform clouds are seeded with dry ice at the rate of about 1 lb. per mile. The magnitude of the effect is dependent on the stability of the air, the thickness of the cloud, and whether or not the air flow into the region is of a convergent nature. If convergence is present, it is quite possible that the "triggering effect" produced by seeding with dry ice would result in a self-perpetuating storm spreading out from the area originally modified. Similarly, cumulus clouds may be profoundly modified with dry ice whenever the cloud contains supercooled water droplets.

Laboratory experiments indicate that at least 10^{16} ice nuclei may be produced from 1 g. of dry ice when it is introduced into a supercooled cloud. If all of these particles could be supplied with moisture and allowed to grow to a small

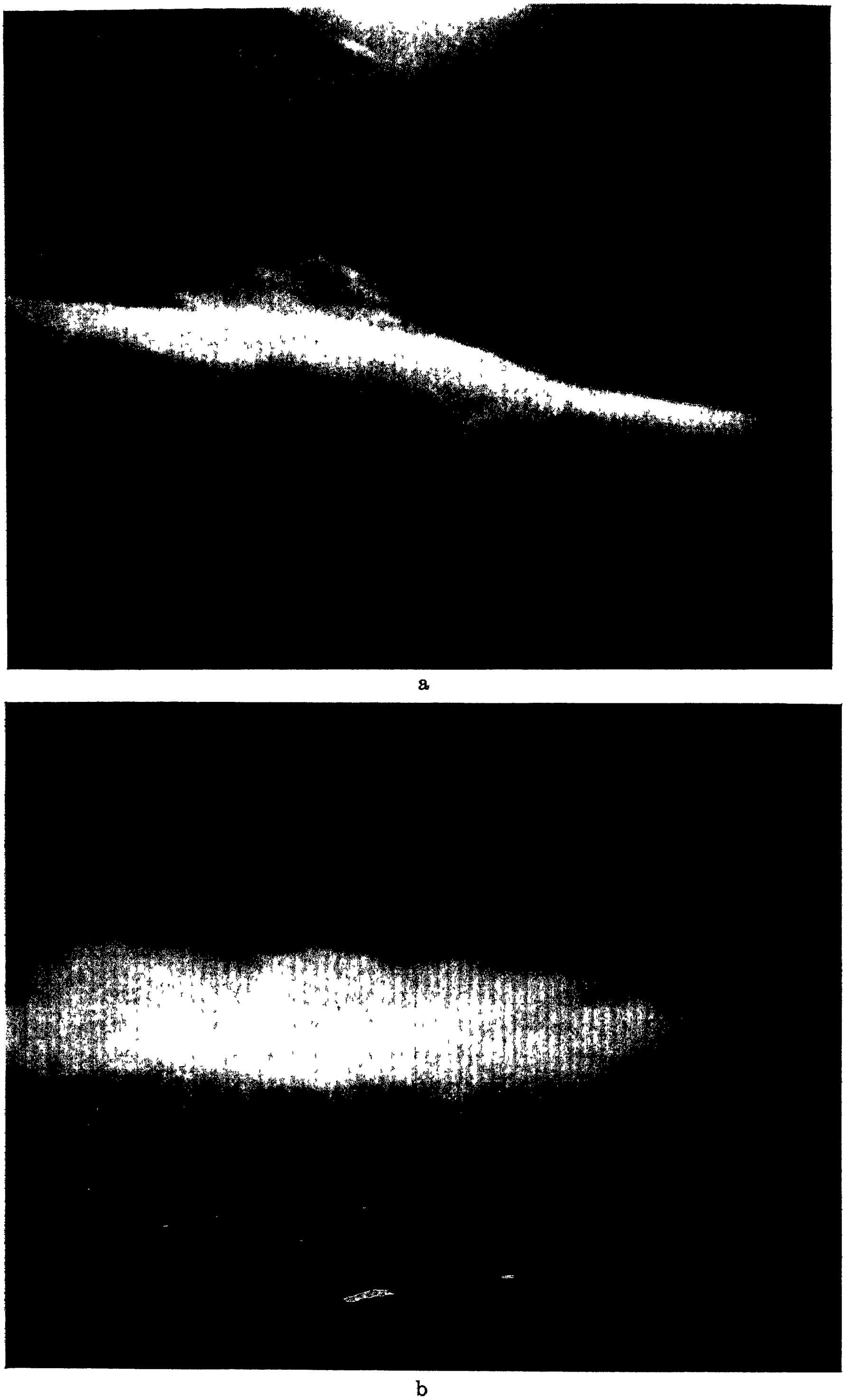


FIG. 20. View of first experiment in natural clouds above Mt. Greylock, Massachusetts (a) before seeding; (b) 5 min. after seeding with dry ice.

snow crystal, some 300,000 tons of snow would result. It is, of course, impossible to achieve such results under natural conditions, since there is not enough moisture available in the area that can be affected. The important fact to recognize in this respect is that by using dry ice it is quite feasible to produce nuclei on the tremendous scale that is characteristic of natural weather phenomena. For example, the number of snow crystals required to produce 1 ft. of snow over an area of 10,000 square miles is approximately 1×10^{18}. That number of ice nuclei could easily be produced with 2 lb. of dry ice. Since it would not be feasible to spread 2 lb. of dry ice over such a large area, it would, however, be practical to use 200 lb. at the rate of ½ lb. per mile and exert a large influence over an area of 10,000 square miles.

Under actual seeding operations, it has been found that about 1 lb. of granulated dry ice will convert 5 cubic miles of supercooled stratus clouds to snow in about 30 min. when the dry ice is introduced from a 2-in. pipe in a plane flying at the top of a cloud having a vertical thickness of a mile. Dry ice pellets in the size range 1–10 mm. in diameter are used so that the crystals are produced within a minute or so along a plane extending from top to bottom of the cloud. The heat generated in the crystallization process ranges from 0.5 to 1°C. in such a cloud, causing turbulence which spreads the crystals upward and outward. Thus it is a common observation to see the affected area expand as shown in figure 21 at a rate of nearly 5 miles an hour. Figure 22 is a photograph of such an area.

So many crystals are produced when dry ice is dispensed at the rate of 1 lb. per mile that the affected region often shows a persistent, stable snow crystal cloud remaining in the modified region after all of the adjoining clouds have disappeared by precipitation. Such an effect is recognized as "overseeding" but is only experienced with very stable clouds, thin clouds, or under conditions free of convergence.

Such effects would not be expected with cumulus clouds unless they were dissipating or when large quantities of dry ice are used.

Some evidence has been seen that extensive cumulus activity may be generated in thick and unstable stratiform cloud systems by certain seeding techniques. Studies are now under way by our group to explore these possibilities.

SOME POSSIBILITIES AND LIMITATIONS IN MODIFYING CLOUDS

There are, of course, definite limitations to the effects which may be expected when supercooled cloud systems are artificially modified with a material such as dry ice.

If any appreciable amounts of precipitation are to result from such activities, the atmospheric conditions must be such that it is likely that snow or rain would occur from natural processes. The method does, however, present the definite possibility of "triggering off" a storm sooner than it would develop if the natural sequence prevailed.

It must be remembered that there is actually only a relatively small amount of condensed water in a cloud. Thus, if all of the moisture in an ordinary stratus cloud having a vertical thickness of a mile were to be precipitated, only about 2 mm. of water would result. Towering cumulus of the congestus type may contain from 1 to 3 cm. of precipitable water, depending on vertical height and liquid water content.

While it is not possible to produce ice crystals in clouds using the dry ice method unless they are *cold* clouds, with temperatures colder than 0°C., it is obvious that if a *cool* cloud were modified in the supercooled region, the resulting snow crystals would melt and become rain at a short distance below the freezing level, thus providing water droplets which would lead to the formation of rain. If Langmuir's chain reaction theory is correct, this effect would quickly lead to a rapid development of heavy rain in large cumulus clouds.

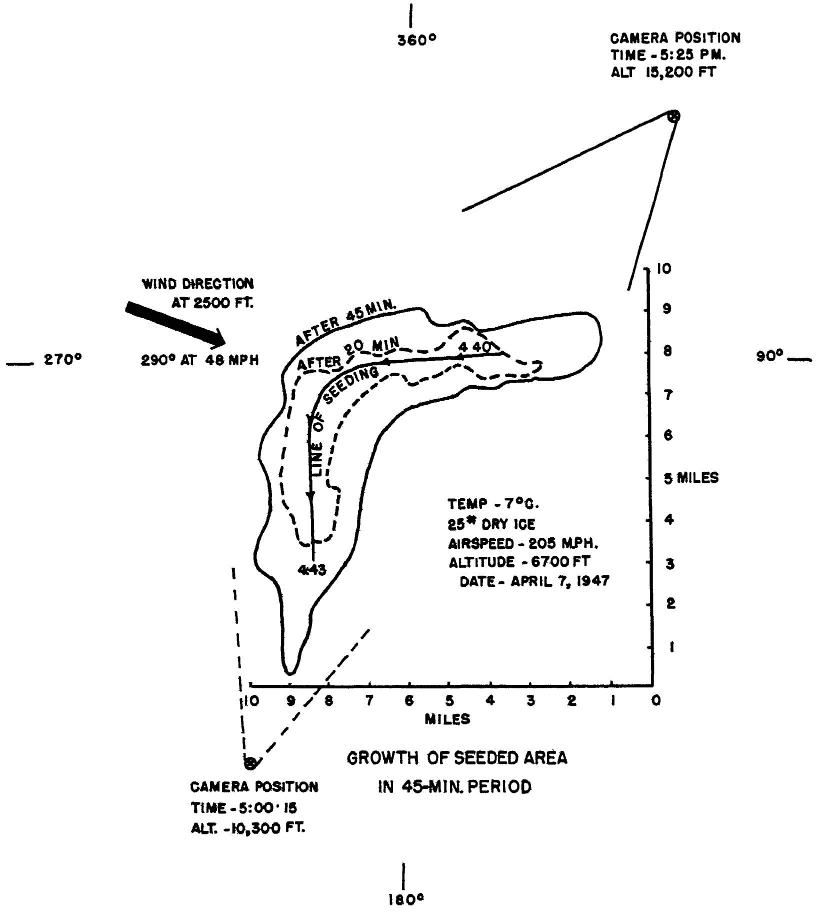


FIG. 21. Rate of growth of area in stratus cloud seeded by Project Cirrus

The tremendous quantities of ice crystals which may be produced with dry ice in cold clouds raises the distinct probability that, if correctly introduced, it should be possible to eliminate dangerous hailstorms. With a large concentration of sublimation nuclei present in the region where hail might form, the competition between particles for the available moisture would be so keen that no particles could grow large. It would, indeed, be foolish to think that such results could be obtained immediately. Much careful quantitative research experiments must be conducted to hope for eventual success in this field. In a similar manner, violent thunderstorms may eventually be subdued by preventing the development of any appreciable thickness of supercooled cloud in the cold portion of the cool cumulus congestus cloud.

In fact, any type of storm system which contains appreciable amounts of supercooled clouds may be profoundly modified by intelligent seeding with dry ice.

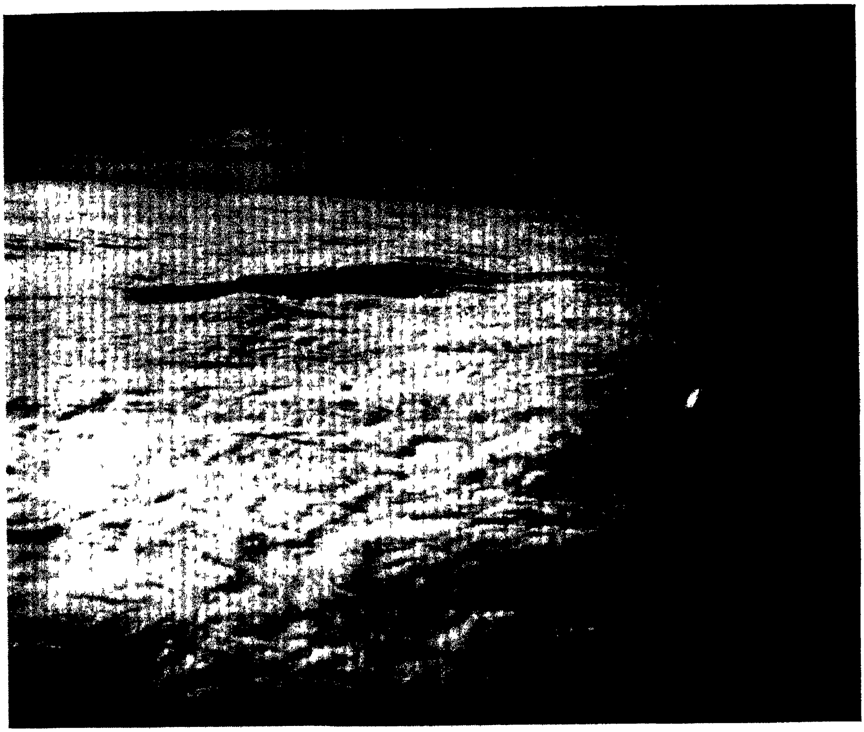


Fig. 22. Photograph of modified stratus cloud seeded at rate of 1 lb. per mile. The modified area which appears dark is 15 miles long and 2 miles wide.

Supercooled ground fogs may also be modified by following a proper seeding procedure. There is a likelihood of "overseeding" with ground fog which not only reduces visibility but also makes the fog much more persistent. Proper amounts used will, however, clear the fog quite rapidly.

There is also a real possibility that the aircraft icing problem may be affected to a considerable degree by modifying supercooled clouds in the vicinity of airports and along heavily travelled air lanes. There is also a distinct possibility that the local area ahead of a flying aircraft may be modified with projectiles carrying and dispensing seeding media, such as dry ice from liquid carbon dioxide. In this way, it may be possible that the icing hazard could be reduced or eliminated as the plane reached the treated area.

One of the immediate results which is proving to be of much value in the physical studies of clouds is the fact that the dry ice seeding process permits the marking of clouds in a unique way. By photographing the subsequent appearance of the modified cloud, many interesting features may be measured in a quantitative manner by special photographic techniques and procedures under development by our Project Cirrus operation.

The science of experimental meteorology is just beginning. The next few years will show rapid advances in this fascinating field. The results will surely aid meteorologists and others in reaching a better understanding and appreciation of weather phenomena.

REFERENCES

(1) Bentley, W. A., and Humphreys, W. J.: *Snow Crystals*. McGraw-Hill Book Company, Inc., New York (1931).
(2) Dobrowolski, A. B.: *Historja Naturalna Lodu*. Warsaw (1923).
(3) Findeisen, W.: Meteorol. Z. **55**, 121–33 (1938).
(4) First Quarterly Progress Report, Meteorological Research: obtainable from Office of Technical Services, Department of Commerce, Washington 25, D. C.
(5) Hellmann, G.: *Schnee Krystalle*. Berlin (1893).
(6) Langmuir, I.: J. Meteorol. **5**, 175–92 (1948).
(7) Nakaya, U., *et al.*: J. Faculty Sci. Hokkaido Imp. Univ., Ser. II, **2**, No. 1 (1936).
(8) Nakaya, U., and Sekido, Y.: J. Faculty Sci. Hokkaido Imp. Univ., Ser. II, **1**, No. 9 (1935).
(9) Schaefer, V. J.: Science **93**, 239–40 (1941).
(10) Schaefer, V. J.: Nature **149**, 81 (1942).
(11) Schaefer, V. J.: "Final Report on Icing Research up to July 1, 1945," General Electric Research Laboratory. A.T.S.C. Contract W-33-038-AC-9151.
(12) Schaefer, V. J.: Science **104**, 457–9 (1946).
(13) Schaefer, V. J.: Trans. Am. Geophys. Union **28**, 587 (1947).
(14) Schaefer, V. J.: J. Inst. Navigation **1**, 172–4 (1947).
(15) Schaefer, V. J.: Weatherwise **1**, 3–5 (1948).
(16) Schaefer, V. J.: Bull. Am. Meteorol. Soc. **29**, 175–82 (1942).
(17) Schaefer, V. J.: "The Detection of Ice Nuclei in the Free Atmosphere," to be published in the Journal of the American Meteorological Society (1949).
(18) Schaefer, V. J.: Trans. Am. Geophys. Union **29**, 492–8 (1948).
(19) Vonnegut, B.: J. Applied Phys. **18**, 593–5 (1947).
(20) Vonnegut, B.: Science **107**, 621–2 (1948).

ON THE NATURE AND SIZE OF PARTICLES IN HAZE, FOG, AND STRATUS OF THE LOS ANGELES REGION[1, 2]

M. NEIBURGER AND M. G. WURTELE

Department of Meteorology, University of California at Los Angeles, Los Angeles, California

Received October 18, 1948

Measurements of relative humidity in fog and stratus cloud frequently give values far below 100 per cent, particularly in industrial areas. Drop radii measurements in the cloud show most frequent values around 7 μ. The occurrence of drops in equilibrium at a relative humidity of 90 per cent would require solutions of concentrations of 17 per cent for sodium chloride and 23 per cent for sulfuric acid. Particularly in the latter case, i.e., for combustion nuclei, this high concentration seems unlikely. The hygroscopic nuclei would be about one-half the size of the drop.

Visibility measurements, in conjunction with the humidity, give a means of estimating the relative contribution of hygroscopic and non-hygroscopic particles in obstructing vision, on the assumption that the number of hygroscopic particles is constant. The observations indicate that the latter assumption is not valid. Using a relationship between liquid water content and visibility shown by Radford to fit observed data, a method is derived for determining the number and size of the hygroscopic particles, the value of the extinction coefficient due to non-hygroscopic particles, and the mass of the nucleus. Using the observations of humidity and visibility at Los Angeles Airport, and assuming that the nuclei are sodium chloride, these quantities are determined. The drop radii vary from 8 microns at 98 per cent humidity to 3 microns at 67.5 per cent, and the number decreases from 24 to 7 per cubic centimeter.

I. INTRODUCTION

In the California Stratus Investigation of 1944 (11) it was noted that the maximum relative humidity recorded in the cloud at different stations in the vicinity of Los Angeles was frequently less than 100 per cent. While at Santa Maria, a rural station, the recorded humidity reached 100 per cent in most of the soundings through stratus, there were a number of soundings in which it did not. At U.C.L.A., in the western part of Los Angeles, and in Pasadena, to the east of it, the highest humidity recorded in the cloud was most frequently about 90 per cent. It was suggested, without careful consideration, that "a plausible explanation of this difference lies in the fact that there is little industrial activity in the vicinity of Santa Maria, but in the Los Angeles Basin there are a great number of factories injecting combustion products and other forms of hygroscopic nuclei into the air." When this explanation is examined, however, serious doubts are raised as to its validity, as we shall show below.

---

[1] Presented at the Symposium on Aerosols which was held under the joint auspices of the Division of Physical and Inorganic Chemistry and the Division of Colloid Chemistry at the 113th National Meeting of the American Chemical Society, Chicago, Illinois, April 22, 1948.

[2] U. C. L. A. Department of Meteorology, Papers in Meteorology No. 8. This study was carried out as part of a project supported by the Office of Naval Research.

The idea that combustion nuclei may lower considerably the humidity at which condensation occurs is presented quite generally in the literature. Thus Shaw (17) says, "It is hardly safe to assert that there is any limit of humidity below which clouds of water particles could not be formed in the air of industrial cities." On the other hand Wright (22) attempted to evaluate the contribution of sea-salt nuclei, combustion nuclei, and non-hygroscopic particles from average visibilities at coastal and inland stations in Britain in summer and winter, and concluded that "Water fogs are due to sea-salt nuclei and if supersaturation is attained the fogs may thicken considerably owing to the deposition of water on the sea-salt nuclei which consequently become very much larger. Such deposition does not occur on the combustion nuclei if sea-salt nuclei are present."

II. ON THE NATURE AND SIZE OF NUCLEI

The three factors altering the equilibrium vapor pressure over drops from that over a plane pure water surface at the same temperature are curvature of surface, electric charge, and dissolved substances. Of these factors the first tends to raise the equilibrium vapor pressure, and the other two to reduce it.

The curvature effect is given by the formula

$$\ln \frac{e_r}{E} = \frac{2\gamma}{\rho R T r} \tag{1}$$

where e_r and E are, respectively, the equilibrium vapor pressures over a drop of radius r and a plane liquid surface at absolute temperature T, ρ is the density, γ is the surface tension of the liquid, and R is the gas constant. For pure water at 15°C. this formula becomes

$$\ln \frac{e_r}{E} = \frac{1.11 \times 10^{-3}}{r}$$

where r is expressed in microns. A 1-micron drop would thus have an equilibrium vapor pressure 1.001 times that for a plane water surface, i.e., would require a "supersaturation" of 100.1 per cent relative humidity in order to be in equilibrium. For larger drops the amount of supersaturation required is still less.

Figure 1 shows the frequency distribution of drop sizes measured by the sooted-slide method at several levels in a stratus cloud (12). The most frequent radius at all levels except near the cloud base is about 7 microns, and smaller drops were much less frequent. The preponderance of drops of one size is to be expected for drops formed by condensation, since larger drops will grow at the expense of smaller ones, other conditions being the same. The example shown in the figure is typical of the several cases in which measurements were made in California stratus. Thus for this cloud the drops are so large that the effect of curvature may be neglected.

That the effect of electric charge is also negligible may be seen by computing the charge required to offset the effect of curvature. If the drop has a charge of q electrostatic units, equation 1 becomes

$$\rho R T \ln \frac{e_r}{E} = \frac{2\gamma}{r} - \frac{q^2}{8\pi r^4}$$

To offset the curvature effect in a 1-micron drop would require 1.3×10^5 times the charge of one electron. For a 7-micron drop it would require 2.36×10^6 electronic charges. It is unlikely that such large charges ever occur. The effect of solutes (Raoult's law) is the only one remaining which might explain the occurrence of drops at less than 100 per cent relative humidity.

For solutions the reduction in vapor pressure is given by

$$\frac{E}{e_s} = 1 - CM \tag{2}$$

where e_s is the equilibrium vapor pressure for pure water, M is the molar concentration of the solution, and C is a factor which depends on the concentration, temperature, and nature of the solute. $E/e_s = H$ is the relative humidity with respect to a plane pure water surface. If c is the mass concentration, $c = mM$, where m is the molecular weight. Then

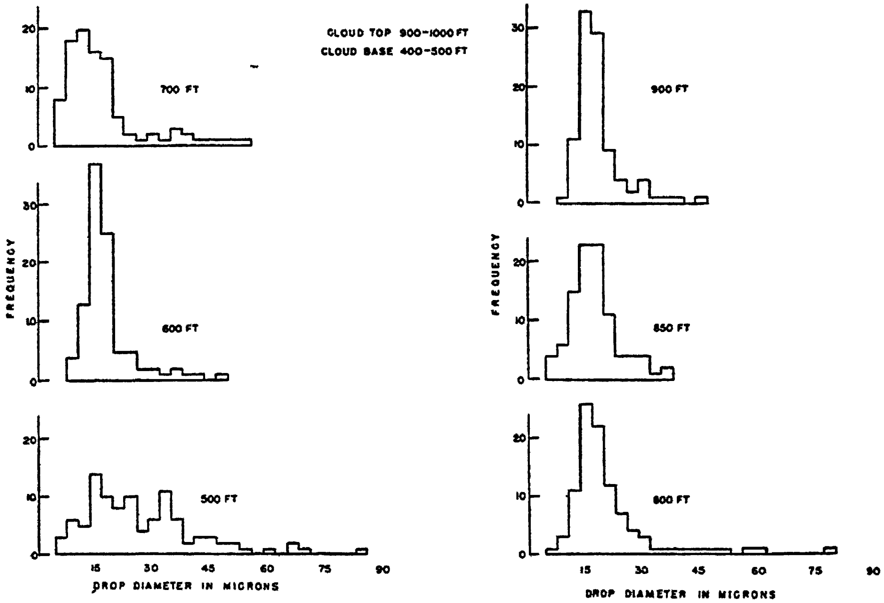


FIG. 1. Drop-size distribution in one case of California stratus, as measured by the sooted-slide method.

$$H = 1 - cC/m \tag{2'}$$

The two types of solute which have been most frequently suggested as condensation nuclei under atmospheric conditions are sea salt, of which sodium chloride is the principal constituent, and combustion products, of which sulfuric acid is regarded as the most hygroscopic. For M of the order of 0.002 at 20°C., $C = 34$ for sodium chloride and $C = 43$ for sulfuric acid (6).

The concentration of solution in equilibrium at 90 per cent relative humidity is found by substituting 0.9 for E/e_s in equation 2. We get

$$c = 0.1\, m/C$$

which for sodium chloride and sulfuric acid gives the concentration of 17 per cent and 23 per cent, respectively. These high concentrations seem surprising. If the fog drops consist of acid of this concentration, the corrosive effects would have been noticed. Since they have not, it is reasonable to reject the existence of such high concentration of acid.

If we characterize the size of the nucleus by the radius r_s of a sphere of equal mass, we can compute the ratio of nucleus to drop size which will result in this concentration. The mass of the solute in one drop is given by the two expressions

$$m_s = \tfrac{4}{3}\pi r^3 \rho C = \tfrac{4}{3}\pi r_s^3 \rho_s$$

Thus

$$\frac{r_s}{r} \qquad \sqrt{\frac{\rho C}{\rho_s}} \tag{3}$$

The following table gives the values of this ratio for sodium chloride and sulfuric acid for the concentrations found above, and the corresponding nuclear size and mass for 7-micron drops.

SOLUTE	ρ_s	C	ρ	r_s/r	$r_s(\mu)$	m_s
						grams
NaCl....	2.17	0.171	1.12	0.44	3.1	2.75×10^{-10}
H_2SO_4.	1.83	0.228	1.17	0.53	3.7	3.83×10^{-10}

Since John Aitken's invention of the dust counter there have been hundreds of studies of the properties of "nuclei" (for summaries see Landsberg (9) and Simpson (18)). Most of these, however, have been concerned with those whose size can be determined from their ionic mobility, i.e., of radius of the order 10^{-6} cm. or less (see, for example, Israel and Schulz (7) or Nolan and Guerrini (13)). The properties of larger nuclei are, paradoxically, more difficult to determine and require more circuitous methods. Köhler (8), by measurements of corona and chemical analysis of frost deposits obtained when supercooled fog passed over objects, concluded that the nuclei were particles of sea salt with "radius" of the order 10^{-5} cm. Similar results were obtained by Wright (21, 22, 23) by an analysis of observations of humidity and visibility. Simpson (19) raised objections to Wright's results. With respect to Köhler's measurements, there is considerable question of the applicability of the formula he used relating corona diameter and drop size to the range of sizes involved. In table 1 are presented the ranges of values reported by these and other observers, and now tentatively accepted.

The measurements of nuclei thus have values considerably smaller than those necessary to maintain equilibrium of drops of solution at 90 per cent relative humidity. In particular, the combustion nuclei have been found to be at largest less than $\frac{1}{30}$ the size required. Nuclei of sea salt have been found to be quite a bit larger than combustion nuclei. We must conclude that combustion nuclei cannot be responsible for the difference in humidity measured in the stratus at Los Angeles and Santa Maria.

The possibility that the low humidity measurements in the Los Angeles area are due to instrumental errors was considered. It was pointed out in the 1944 report that this is rendered doubtful by the fact that the same types of instruments and methods of evaluation were used at all stations. Besides this, surface observations show the existence of fog at humidities considerably below 100 per cent, even at rural stations, with more frequent low humidities in fog in industrial areas. Thus Pick (15, 16) showed that 20 per cent of all thick fogs observed at 0700 at Cardigan, England, in 1929–30 were accompanied by

TABLE 1

Ranges of sizes of nuclei (various sources)

NUCLEUS	MASS	RADIUS
	grams	*cm.*
Combustion	2×10^{-16} to 5×10^{-15}	10^{-6} to 10^{-5}
Sea salt	3.5×10^{-16} to 2.8×10^{-13}	3×10^{-6} to 4×10^{-5}

TABLE 2

Frequency of relative humidities at time of lowest visibility

REPORTED RELATIVE HUMIDITY*	NUMBER OF OBSERVATIONS AT INLAND STATIONS	PER CENT OF TOTAL		NUMBER OF OBSERVATIONS AT MARINE STATIONS	PER CENT OF TOTAL	
per cent						
100	12	20.7	41.4	23	29.9	80.5
99	0	0		2	2.6	
98	0	0		2	2.6	
97	12	20.7		26	33.7	
96	0	0		4	5.2	
95	0	0		5	6.5	
94	14	24.2	51.7	4	5.2	19.5
93	0	0		2	2.6	
92	0	0		1	1.3	
91	2	3.4		2	2.6	
90	12	20.7		6	7.8	
89	2	3.4		0	0	
88	0	0	6.9	0	0	0
87	0	0		0	0	
86	3	5.2		0	0	
85	0	0		0	0	
84	0	0		0	0	
83	1	1.7		0	0	
Total	58	100.0		77	100.0	
Average relative humidity	94 per cent			97 per cent		

* The curious concentration at the humidities 97 per cent and 94 per cent in columns two and four may be due to the fact that they may be read from the psychrometric tables without interpolation.

relative humidities of 90 per cent to 93 per cent, and that the majority of all fogs occurred in unsaturated air. He also quoted observations of thick and dense fogs at sea with humidities in the 80's.

That the difference between purely coastal and industrial areas exists with respect to surface observations of fog and humidity in the Los Angeles region is shown by comparison of stations relatively free from industrial influence with those so affected. The two groups of stations considered were CIT, Ontario, and San Bernardino as inland locations and Los Angeles Municipal Airport, Oceanside, and Santa Barbara on the coast. Table 2 lists the frequency of observed (wet-and-dry-bulb psychrometer) relative humidities at the time of lowest visibility on each morning during a period of several months, provided this visibility was 1 mile or less.

TABLE 3

Frequency of minimum visibilities reported for data in table 2

REPORTED VISIBILITY	NUMBER OF OBSERVATIONS AT INLAND STATIONS	PER CENT OF TOTAL		NUMBER OF OBSERVATIONS AT MARINE STATIONS	PER CENT OF TOTAL	
miles						
0	5	8.6	34.4	16	20.8	41.6
1/8	8	13.8		7	9.1	
1/5	2	3.4		2	2.6	
1/4	5	8.6		7	9.1	
1/2	12	20.7	65.6	12	15.6	58.4
3/4	10	17.3		14	18.2	
1	16	27.6		19	24.6	
Total.......	58	100.0		77	100.0	
Average visibility.......	9/16 miles			7/16 miles		

This distribution must be compared with a similar one for the visibilities, with intervals determined by standard airway observation practice, shown in table 3.

While the visibilities at marine stations were somewhat lower than at inland stations for the cases considered, the difference is not nearly so marked as that in the humidities. We see that while only 19.5 per cent of the cases of low visibility had humidities below 95 per cent at marine stations, the corresponding value was 58.6 per cent at inland stations.

With the hygroscopic action of combustion nuclei rejected as a factor in explaining the more frequent existence of cloud drops at low humidities, and further evidence that the phenomenon is real, we are led to seek other explanations. One hypothesis which might be the explanation is that the drops are not in equilibrium, but have not had time to evaporate completely at the time of observation. Findeisen (1, 2) derived an equation for the time of evaporation of drops. For relative humidity 90 per cent, temperature 5°C., and pressure

900 mb., his formula is

$$t = 0.053r^2$$

where r is in microns and t in seconds. Seven-micron drops would thus take 2.5 sec. to evaporate under these conditions, and from this Findeisen concludes that "the probability of small drops being present in air only slightly below saturation is practically zero."

III. VISIBILITY AND PARTICLE SIZE

The preceding material points up the desirability of direct measurement of the nature and size of condensation nuclei. As was pointed out above, such direct measurement is extremely difficult. However, Wright's approach may be applied to the determination of the drop and nucleus size, if assumptions are made as to the nature and number of active hygroscopic particles. In this way hypotheses regarding their nature and number can be tested, using only the regular Weather Bureau observations of visibility and humidity.

The relationship between the visibility (visual range) and the suspended particles which limit it is given by Koschmieder's formula (10)

$$V = \frac{1}{\sigma} \ln \frac{1}{\epsilon} = \frac{3.91}{\sigma} \tag{4}$$

where V is the visibility in centimeters, σ is the extinction coefficient due to suspended matter, and ϵ, the threshold of perceptible contrast between the apparent brightness of an object and the brightness of the background, is taken to be 0.02. The extinction coefficient σ is contributed to by *(1)* scattering by molecules, *(2)* scattering and absorption by hygroscopic particles and their associated water drops, and *(3)* scattering and absorption by non-hygroscopic particles of dust, smoke, etc. We may write

$$\sigma = \sigma_M + \sigma_N + \sigma_P$$

where σ_M, σ_N, and σ_P are the separate extinction coefficients for processes *(1)*, *(2)*, and *(3)*, respectively.

For average wave length of light and normal temperature and pressure Hulburt (5) computed $\sigma_M = 1.6 \times 10^{-7}$ cm.$^{-1}$ For water drops σ_N is given by Stratton and Houghton (20) to be (see appendix)

$$\sigma_N = 2\pi Nkr^2 \tag{5}$$

where N is the number of drops per cubic centimeter and πr^2 the area of each, and k is a factor depending on the ratio of r to the wave length of the light. For r several times the wave length, k departs only slightly from 1. A similar expression would hold for drops consisting of solutions, but with a different value of k, and presumably also for σ_P, with a still different factor. We shall assume that in both these cases k is also near unity for the wave lengths in which we are interested.

Equation 4 may now be written

$$\sigma_P + 2\pi N r^2 = \frac{3.91}{V} - 1.6 \times 10^{-7} \qquad (6)$$

For values of visibility up to 25 km. the second term on the right is negligible. For convenience we shall absorb it in σ_P. The visibility is observed, and the unknowns in this equation are σ_P, N, and r. If successive measurements are made in the same air at different humidities, but away from sources of pollution, r and N may change, but σ_P should remain constant. The change in drop radius will be in accord with equation 2′. Since the concentration in the drop is given by

$$c = 3m_s/4\pi r^3 \rho$$

we obtain from equation 2′

$$r = K[m_s/(1 - H)]^{1/3} \qquad (7)$$

where $K = (3C/4\pi\rho m)^{1/3}$ is a parameter of the nuclear substance which has a variation of the order of 1 per cent with temperature and concentration. For sodium chloride and sulfuric acid it has the values 0.50 and 0.45, respectively.

Substituting for r from equation 7 in equation 6 we obtain

$$\sigma_P + 2\pi N K^2[m_s/(1 - H)]^{2/3} = 3.91/V \qquad (8)$$

If m_s and N are constants with changing humidity, as well as σ_P, the equation gives the relationship between humidity and visibility in the form

$$(1 - H)^{-2/3} = AV^{-1} + B$$

where A is a parameter depending on N, K, and m_s, and B depends on σ_P. If a value of σ_P is assumed, the curve representing this equation is determined by one observation; other observations may be used to check it. The value $\sigma_P = 0$ leads to the familiar curves of Wright (figure 2).

For the purpose of testing this relationship hourly observations of visibility and humidity at the Los Angeles Municipal Airport were tabulated. Only days on which there were periods with visibilities 2 miles or less were used, and precautions were taken to eliminate effects which would alter σ_P. For instance, if a sea breeze brought about a change of air mass over the station, or if advection of smoke was evident, the observations were not used. Thus σ_P might change from day to day, but not during the series of hours in which the variation of visibility with humidity on a single day was tabulated, insofar as it was possible to select data so as to avoid it. The visibilities were grouped and averaged by 5 per cent intervals of humidity. The results are shown by the broken line in figure 3.

Using the point on the curve for 85 per cent relative humidity to determine A, the theoretical curves representing equation 8 were computed and drawn for two values of σ_P. It is seen that while these curves, particularly the one for $\sigma_P = 2 \times 10^{-7}$, agree fairly well with the observations for high humidities, they diverge widely from them for lower humidity values. To get a closer fit it would be necessary to assume σ_P negative, a physically unacceptable assumption.

Thus the hypotheses of σ_P and N constant must be reconsidered. Since the data were selected with an attempt to keep σ_P constant, and since the constant N requires σ_P to be negative, it is logical to abandon the second hypothesis first. No alternative suggestion has been made regarding the variation of N, nor does one arise from physical reasoning. However, W. H. Radford's compilation of measurements of the content of liquid water in fogs (4) gives a clue to the behavior of drops under varying humidities. He plotted observations of visibility against observations of liquid content, both on a logarithmic scale, and found the curve of best fit to be a straight line (figure 4). Thus if w is the liquid content in grams per cubic centimeter of air, there is empirical evidence that $\ln V = \ln a + b \ln w$ or, more conveniently,

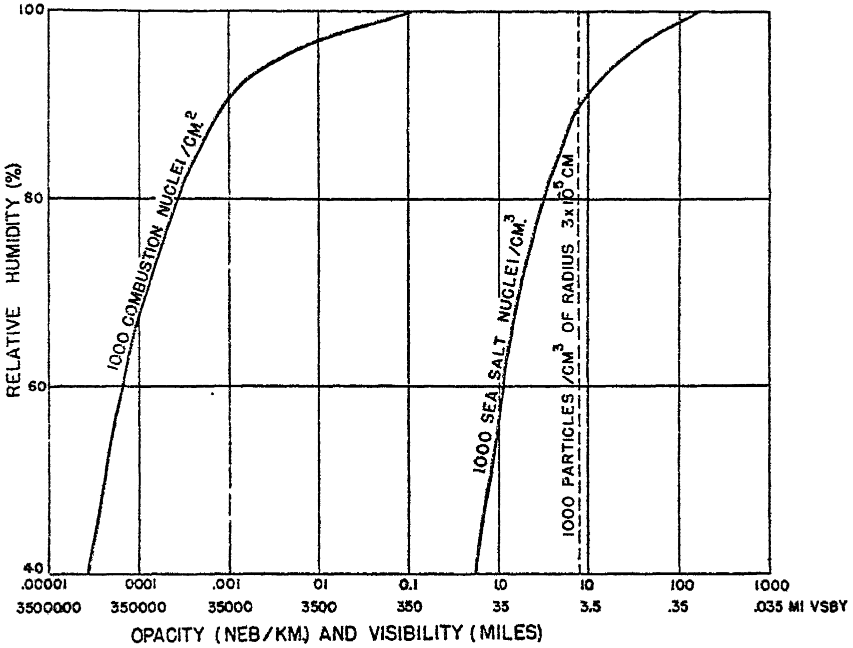


FIG. 2. Relation between relative humidity and visibility, assuming N constant and $\sigma_P = 0$ (after Wright).

$$V = aw^b \tag{9}$$

Radford does not give the constants a and b, but from his graph it is found that $a = 0.032$ and $b = 1.43$.

This empirical relationship may be incorporated into our system of equations in place of the assumption N = constant. The liquid content is

$$w = \tfrac{4}{3}\pi\rho r^3 N \tag{10}$$

and if N is eliminated between this equation and equation 6 we have

$$\sigma_P + \frac{3w}{2\rho r} = \frac{3.91}{V} \tag{11}$$

or, from equation 9

$$\sigma_P + \frac{3}{2\rho r}\left(\frac{V}{a}\right)^{1/b} = \frac{3.91}{V}$$

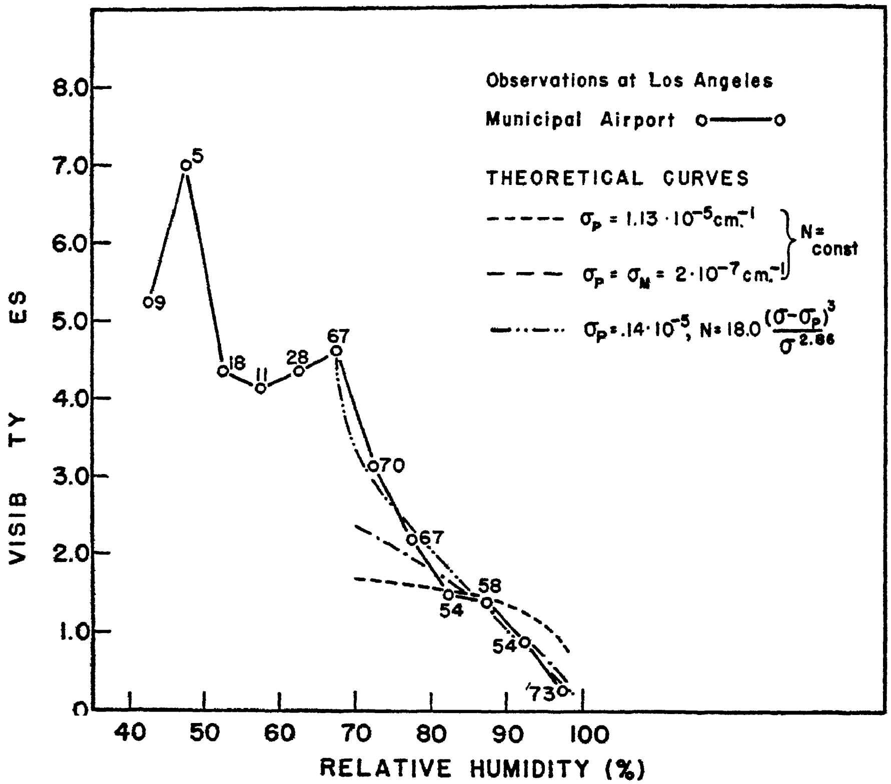


FIG. 3. Observed and computed relations between relative humidity and visibility

Using equation 7, we may write

$$\frac{r_1}{r} = \left(\frac{1 - H}{1 - H_1}\right)^{1/3} \tag{12}$$

where r_1 is the radius at humidity H_1, and r that at H. Thus our equation becomes

$$\sigma_P + \frac{3}{2\rho r_1}\left(\frac{V}{a}\right)^{1/b}\left(\frac{1 - H}{1 - H_1}\right)^{1/3} = \frac{3.91}{V} \tag{13}$$

This gives the relation between humidity and visibility on the basis of our new assumption, giving an implicit relationship between V and H, with σ_P,

r_1, and H_1 as parameters. By choosing two points on the observed curve, H_1 and H_2, and using the observed visibilities V_1 and V_2, the remaining quantities σ_P and r_1 can be determined. By choosing 85 per cent and 67.5 per cent, it was found that $\sigma_P = 0.14 \times 10^{-5}$ cm.$^{-1}$, or 0.14 km.$^{-1}$, and $r_1 = 4.08 \times 10^{-4}$ cm. = 4.08 μ. The curve representing the equation with these values of the parameters is shown in figure 3 also. It comes very close to the observed values, not only at the two points for which it was fitted but in its entire range.

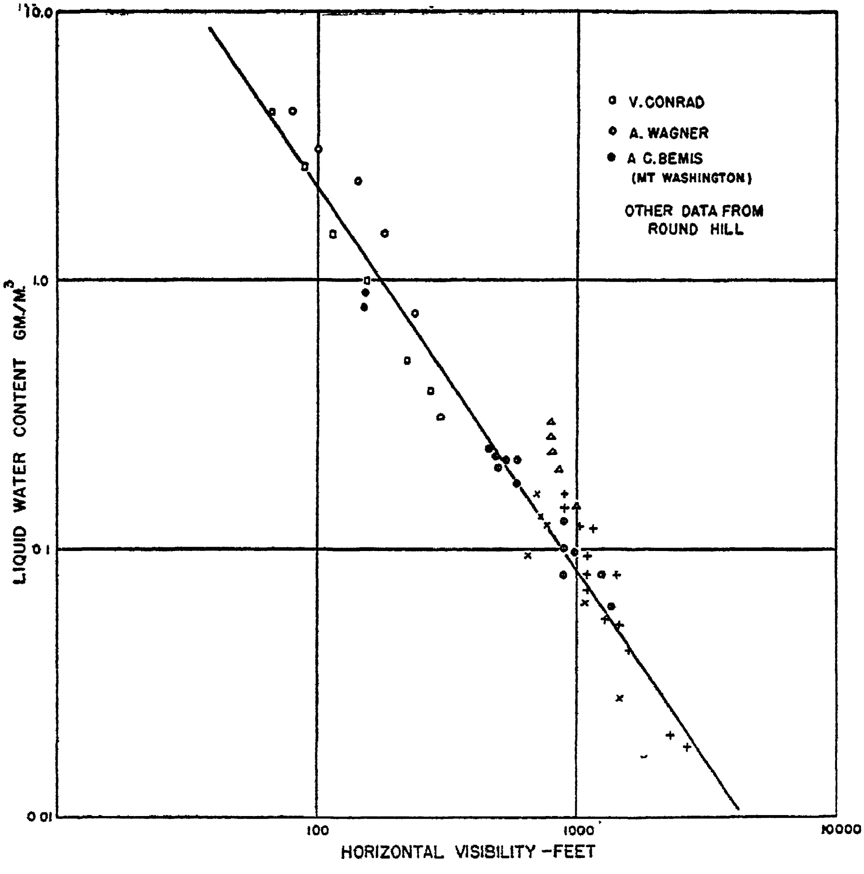


FIG. 4. Relation between liquid water content and visibility, according to summary of observations by Radford.

Using r_1 as determined from equation 13 the values of r may be evaluated from equation 12. The mass of solute m_s is determined from equation 7. The value of K for sodium chloride was used in this connection, and whenever else the solute parameters enter. To evaluate N, we use equations 9 and 10, inserting the values of r computed for the observed values V. The values of these quantities at the different humidities are presented in table 4, along with the computed and observed visibilities.

It is interesting to note that the drop radius determined in this way for fog at

98 per cent relative humidity, 8 μ, is almost exactly the modal radius found for drops in stratus by direct sampling. Even at relatively low humidities the drops remain as large as 3 microns, and in fact the "radius" of the solid nucleus is computed to be 2 microns.

The decrease of the number of drops from 23.5 per cubic centimeter at 98 per cent to 7 per cubic centimeter at 67.5 per cent is not readily explained physically. This, too, corresponds to a phenomenon in the stratus observations. In those observations it was found that while the drop size was constant with height, the liquid content increased linearly from base to top. This means that the number of drops must be greater at the top of the cloud than the bottom. The temperature is lower and humidity presumably higher at the top than the bottom.

TABLE 4

Mean values of various quantities pertaining to suspensoids in the atmosphere (Los Angeles Municipal Airport)

H	$10^5\ \sigma$ cm.$^{-1}$	V	V (observed)	N cm.$^{-3}$	$10^4\ r$
		miles	*miles*		*cm.*
0.98	9.4	0.26	0.2	23.5	7.98
0.965	5.9	0.41	0.4	21.5	6.62
0.95	4.4	0.56	0.6	20.1	5.90
0.90	2.4	1.02	1.1	17.0	4.66
0.85	1.6	1.53	1.5	14.6	4.08
0.80	1.2	2.04	1.8	12.7	3.71
0.75	0.93	2.63	2.6	10.9	3.44
0.70	0.71	3.45	3.8	8.9	3.24
0.675	0.53	4.65	4.6	6.9	3.15
$\sigma_P = 0.14 \times 10^{-5}$ cm.$^{-1}$		$m_s = 8.16 \times 10^{-11}$ g.		$r_s = 2.07 \times 10^{-4}$ cm.	

H = relative humidity.
σ = generalized extinction coefficient.
σ_P = extinction coefficient due to non-hygroscopic particles.
N = number of hygroscopic particles per cubic centimeter.
r = radius of hygroscopic particles.
m_s = mass of solute.

The value of σ_P, 0.14 km.$^{-1}$, is a reasonable one, corresponding to a visibility of 16 miles in the absence of the larger hygroscopic particles. When the marine air containing these large particles is entirely absent in Los Angeles, the visibility is upward of 30 miles in the absence of concentration of industrial pollutants as well. It is seen that in cases of sea haze and fog, if this evaluation is correct, non-hygroscopic particles contribute at most one-fourth to the opacity of the air, and at higher humidities, an insignificant amount. Nothing can be said as to size and number of these particles; if there are 100 per cubic centimeter, the r would have to be about 0.4 micron, while 1000 per cubic centimeter would require, for this value of σ_P, about 0.1 micron radius. It is planned to conduct direct counts to provide additional information regarding these quantities.

In figure 3 it will be seen that the visibility increases rapidly as the humidity decreases to 67.5 per cent, and there becomes approximately constant. This might be considered the value at which the drop becomes crystalline, and thus would give (within 5 per cent, because of the grouping of data) the humidity in equilibrium with a saturated solution of the solute. If the solute were pure sodium chloride, this humidity should be 76 per cent. The experimental determination of the humidity over a saturated solution of sea salt has been carried out by J. S. Owens (14). The experiment was carried out several times at varying temperatures, with the results listed in table 5. Owens' values are thus in complete agreement with the data of figure 3.

IV. CONCLUSION

The methods employed have indicated that it is not combustion nuclei acting alone which produce the difference between humidities in fog and stratus in industrial and rural districts. Large nuclei of sea salt would explain the observed variation of visibility with humidity but cannot explain the difference between rural and industrial areas. The hypothesis that the same number of hygroscopic nuclei are active at all reasonably high humidities must be rejected in favor of one which states that the number increases rather rapidly with humidity. In this connection it should be noted that the methods used did not take account of the contribution of the hygroscopic particles which become inactive to the opacity.

TABLE 5

Relative humidities over saturated solutions of sea salt (after J. S. Owens)

RANGE OF TEMPERATURES	RANGE OF HUMIDITIES	MEAN HUMIDITY
°F.	*per cent*	*per cent*
53–66	60–71	67

TABLE 6

Calculation of percentage error

OBSERVER	RANGE OF $\bar{r}$	RANGE OF PERCENTAGE ERRORS	
		$\frac{\Sigma(r_i - \bar{r})^2}{n\bar{r}^2}$	$1 - \frac{\bar{r}\Sigma r_i^2}{\Sigma r_i^3}$
	microns	*per cent*	*per cent*
Hagemann (3)	6–10	5–17	10–30
Neiburger (figure 1)	7–25	16–32	19–28

If the methods are valid, the data indicate that the drops in fog at the Los Angeles Municipal Airport are of the same size as those found in stratus over the ocean in this region, and in measurements in fog and stratus elsewhere. This would suggest that there is a preferred size for drops in stable air clouds, and an increase in condensation due to further cooling results in more nuclei becoming active, rather than the further growth of existing drops.

Finally it should be emphasized that the numerical results are significant only in order of magnitude, as is shown in the appendix. If greater accuracy is to be attained by the methods employed in this paper, it would seem that further study of drop-size distributions is required. Such a study would necessarily involve consideration of the phenomena discussed in the two preceding paragraphs.

REFERENCES

(1) Findeisen, W.: Meteorol. Z. **56**, 365 (1939).
(2) Findeisen, W.: Meteorol. Z. **56**, 453 (1939).
(3) Hagemann, V.: Gerlands Beitr. Geophys. **46**, 261 (1936).
(4) Houghton, H. G., and Radford, W. H.: *M.I.T. Papers in Physical Oceanography and Meteorology*, Vol. VI, No. 4 (1938).
(5) Hulburt, E. O.: J. Optical Soc. Am. **25**, 125 (1935).
(6) *International Critical Tables*, Vol. III, p. 293. McGraw-Hill Book Company, Inc., New York (1926).
(7) Israël, H., and Schulz, L.: Meteorol. Z. **49**, 226 (1932).
(8) Köhler, H.: Trans. Faraday Soc. **32**, 1152 (1936).
(9) Landsberg, H.: Gerlands Beitr. Geophys., Supplement **3**, 155 (1938).
(10) Middleton, W. E. K.: *Visibility in Meteorology*, 2nd edition, p. 51. The University of Toronto Press, Toronto, Canada (1941).
(11) Neiburger, M., Beer, C. G. P., and Leopold, L. B.: *The California Stratus Investigation of 1944*, U. S. Weather Bureau (1945).
(12) Neiburger, M.: Unpublished data.
(13) Nolan, J. J., and Guerrini, V. H.: Trans. Faraday Soc. **32**, 1178 (1936).
(14) Owens, J. S.: Proc. Roy. Soc. (London) **A110**, 738 (1926).
(15) Pick, W. H.: Quart. J. Roy. Meteorol. Soc. **55**, 305 (1929).
(16) Pick, W. H.: Quart. J. Roy. Meteorol. Soc. **57**, 288 (1931).
(17) Shaw, Sir Napier: *Manual of Meteorology*, Vol. III, p. 330. Cambridge University Press, Cambridge, England (1930).
(18) Simpson, G. C.: Quart. J. Roy. Meteorol. Soc. **67**, 99 (1941).
(19) Simpson, G. C.: Quart. J. Roy. Meteorol. Soc. **67**, 163 (1941).
(20) Stratton, J. A., and Houghton, H. G.: Phys. Rev. **38**, 159 (1931).
(21) Wright, H. L.: Proc. Phys. Soc. (London) **48**, 675 (1936).
(22) Wright, H. L.: Quart. J. Roy. Meteorol. Soc. **65**, 411 (1939).
(23) Wright, H. L.: Quart. J. Roy. Meteorol. Soc. **66**, 66 (1940).

APPENDIX

The assumption made at this point, usually tacit in the literature, is that the proper value of the total surface area of n drops,

$$\pi(r_1^2 + r_2^2 + \cdots r_n^2) = \pi \sum_1^n r_i^2 \tag{14}$$

can be adequately approximated by $\pi n\bar{r}^2$, where $\bar{r}$ is the arithmetic mean of the n radii. The algebraic identity relating these quantities is

$$n\bar{r}^2 - \sum_1^n r_i^2 = -\sum_1^n (r_i - \bar{r})^2$$

so that if there is any dispersion at all there will be an error, which will always have the same sign (negative), with a magnitude equal to n times the variance of the distribution. Calculations of this magnitude from empirical distributions of direct measurements recorded by two observers are presented in column 3 of table 6. It turns out to be rather large, in one instance amounting to 32 per cent.

However, the expression 14 does not enter explicitly into equation 11; rather the substitution from equation 10 results in a different approximation:

$$\sigma_N \propto \sum r_i^2, \qquad w \propto \sum r_i^3$$

Thus

$$\sigma_N = w \frac{\sigma_N}{w} \propto w \frac{\sum r_i^2}{\sum r_i^3}$$

This quantity has been approximated by $w/\bar{r}$. The percentage error is therefore

$$1 - \bar{r} \sum r_i^2$$

Empirically determined values for this error have been tabulated in column 4 of table 6. Again it is large, but does not invalidate the order of magnitude of the results of the paper.

A NOTE ON THE CHANGE IN AVERAGE PARTICLE MASS DURING THE AGING OF AMMONIUM CHLORIDE SMOKES[1]

G. C. LANGSTROTH,[2] T. GILLESPIE,[2] AND R. M. PEARCE

Department of Physics, University of Alberta, Edmonton, Alberta, Canada

Received October 18, 1948

An analysis of particle number and mass concentration data for the aging of ammonium chloride smoke in both still and moving air permits certain deductions concerning the average growth of particles under the experimental conditions used. With still air the average mass per particle increased for at least 5.5 hr., although the rate was smaller in the later stages; with rapidly moving air it increased for a comparatively short initial period and then decreased. The results are quantitatively described by equations involving two loss constants and the coagulation constant, whose values were determined experimentally (1).

In the attempt to obtain detailed information about the changes that occur with time in systems of solids dispersed in air, and about the influence of air motion on them, experiments have been performed on the aging of ammonium chloride smokes. This substance was chosen partly because of ease in handling certain experimental features, but also because it is expected to form a system in which the growth of particles is due primarily to coagulation on collision. The studies were concerned, not with the initial stages of formation of the aerosol, but with the period beginning at 3 min. after complete generation of the smoke (5 min. after the start of generation). Determinations of particle number (n, cc.$^{-1}$) and mass concentration (m, mg./m.^{3}) were made at 30-min. intervals for a period of 5.5 hr. For studies with moving air, a large pierced vane within the chamber was oscillated at a constant rate throughout the experiment. The reading of a non-directional thermocouple anemometer averaged over various positions in the 1.4 m.^{3} chamber was taken as an index of air motion (V, m./min.). A description of apparatus and techniques, and detailed data for n and m have been given previously (1).

It has been shown (1) that the variations of mass concentration and of particle number with time (t, min.) are described adequately by the relations

$$m = m_0 e^{-\alpha t} \tag{1}$$

and

$$\left(\frac{1}{n} + \frac{k}{\beta}\right) = \left(\frac{1}{n_0} + \frac{k}{\beta}\right) e^{\beta t} \tag{2}$$

under the conditions of these experiments, and that an expression of the form of equation 2 may be expected on general grounds when loss of particles to adjacent surfaces is appreciable and the particle size is not too small. In these equations α and β are "loss constants" and k denotes the coagulation constant; m_0 and n_0 denote, respectively, the values of m and n at the start of the observation period ($t = 0$). For "still" air α, β, and k were found to be 3.4×10^{-3} min.$^{-1}$, 3.2×10^{-3} min.$^{-1}$, and 2.2×10^{-8} cc./min., respectively. All three constants were increased by air motion, reaching respective values of 13.4×10^{-3} min.$^{-1}$, 9.8×10^{-3} min.$^{-1}$, and 5.4×10^{-8} cc./min. at an equivalent air flow of $V = 50$ m./min. The value of k for still air was not significantly higher than that expected from the Smoluchowski theory.

---

[1] Presented at the Symposium on Aerosols which was held under the joint auspices of the Division of Physical and Inorganic Chemistry and the Division of Colloid Chemistry at the 113th National Meeting of the American Chemical Society, Chicago, Illinois, April 22, 1948.

[2] Present address: Experimental Station, Suffield, Alberta, Canada.

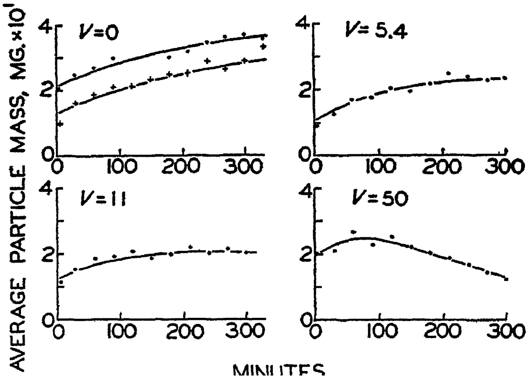


FIG. 1. Typical plots of average particle mass (ordinates, mg. $\times 10^{10}$) against time (abscissae, minutes) for air motion of 0, 5.4, 11, and 50 m. per minute. The points were calculated directly from observed n values and interpolation of determined m values. The curves were calculated from equation 7, using the following experimentally determined (1) constants:

V	μ_0	k	α	β
m./min.	*mg.* $\times 10^{10}$	*cc./min.* $\times 10^8$	*min.*$^{-1} \times 10^3$	*min.*$^{-1} \times 10^3$
0	2.16	2.1	3.2	3.0
0	1.24	2.3	3.6	3.3
5.4	1.08	2.8	6.2	5.5
11.	1.25	2.3	6.8	5.8
50.	1.88	5.8	14.	9.7

This note is concerned with a further consideration of the earlier data (1) in connection with the changes in average particle mass during the aging process.

THE GROWTH OF PARTICLES WITH TIME IN AMMONIUM CHLORIDE SMOKES

As has been shown previously (1) the rates of loss of mass, of particle number by collision with adjacent surfaces, and of particle number by coagulation, are described by the equations:

$$\frac{dm}{dt} = -\alpha m \tag{3}$$

$$\left(\frac{dn}{dt}\right)_l = -\beta n \tag{4}$$

$$\left(\frac{dn}{dt}\right)_c = -kn^2 \tag{5}$$

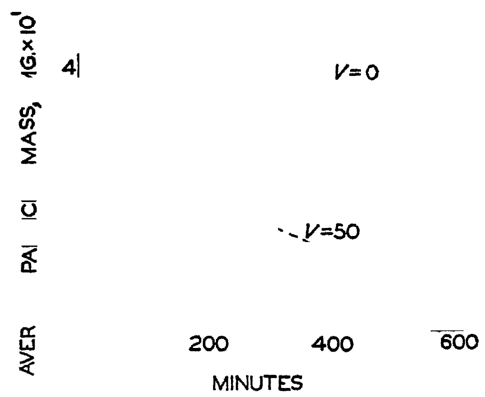


FIG. 2. Curves calculated from equation 7 for "average" smokes (m_0 = 50 mg./m.^{3}, $n_0 = 2.5 \times 10^5$ cc.$^{-1}$) with air motion of 0 and 50 m. per minute. The average values of previously determined (1) constants were used:

V	k	α	β
m./min.	*cc./min.* $\times$ *10*^{8}	*min.*$^{-1}$ $\times$ *10*^{3}	*min.*$^{-1}$ $\times$ *10*^{3}
0	2.2	3.4	3.2
50	5.4	13.4	9.8

Beyond 330 min. the curves are extrapolations.

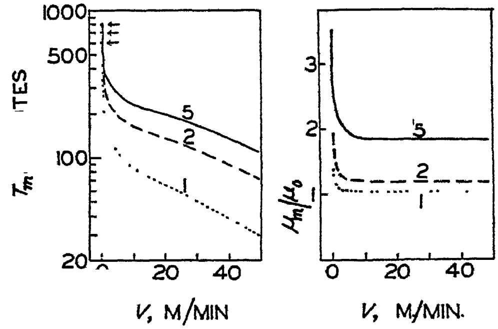


FIG. 3. The dependence on the degree of air motion of T_m and μ_m/μ_0 as calculated from equations 8 and 9. Mean values of the constants for various V values were taken from reference 1. The numbers attached to the curves indicate the n_0 values (cc.$^{-1}$ $\times 10^{-5}$).

Hence the average mass per particle in a hypothetical aerosol for which no surface loss occurs should increase linearly with time according to

$$\mu = \mu_0 + m_0kt \tag{6}$$

and should double in a time $T_2 = 1/n_0k$. Increase in the initial number of particles or in the degree of air motion (i.e., in k) would result in a corresponding decrease in T_2. On the other hand, in a hypothetical non-coagulating aerosol subject to surface loss the average mass per particle should decrease exponentially with time according to

$$\mu = \mu_0 e^{-(\alpha-\beta)t} \tag{6}$$

In the ammonium chloride smokes studied both coagulation and loss occurred, the two mechanisms tending to affect the average particle mass in opposite directions. It follows from equations 1 and 2 that

$$\mu = \frac{m}{n} = \mu_0 e^{-(\alpha-\beta)t} \cdot \left(1 + \frac{n_0 k}{\beta}(1 - e^{-\beta t})\right) \tag{7}$$

The closeness of fit of this expression with individual experimental data is indicated in figure 1 for various degrees of air motion. (In calculating the experimental points it was necessary to interpolate mass concentration data, since samples were not taken simultaneously with those for particle number determinations.) It is noteworthy that according to equation 7 the average mass per particle should reach a maximum value μ_m in a time T_m given respectively by

$$\mu_m = \mu_0 \frac{n_0 k}{\beta} \cdot \frac{\beta}{\alpha - \beta} \cdot e^{-\alpha T_m} \tag{8}$$

$$T_m = \frac{1}{\beta} \ln \left(\frac{n_0 k/\beta}{1 + n_0 k/\beta} \cdot \frac{\alpha}{\alpha - \beta}\right) \tag{9}$$

The trend toward this maximum is seen in the curves of figure 1, particularly in those for rapidly moving air.

The effect of air motion on the average particle mass during aging is illustrated by figure 2. The curves were calculated from equation 7, using average values for α, β, and k as previously determined (1). It is indicated that in rapidly moving air the average particle mass was considerably less than that in still air, except in the early stages of aging. This finding receives support from a comparison of the frequency distribution curves for apparent particle radius as previously obtained (1).

The dependence on air motion of the maximum value attained by the average particle mass and the time taken to attain it is illustrated in figure 3. The curves were calculated from equations 8 and 9, using average values of α, β, and k (1). The calculations represent some extrapolation of the original experimental data which covered a somewhat smaller range of initial particle number. Striking features are (*a*) the drastic reduction of T_m and μ_m by relatively slight air motion, and (*b*) the comparatively low value of μ_m attained.

REFERENCE

(1) LANGSTROTH, G. O., AND GILLESPIE, T.: Can. J. Research **B25**, 455 (1947).

THE LAWS OF DEPOSITION AND THE EFFECTIVENESS OF INSECTICIDAL AEROSOLS[1]

V. K. LA MER

Department of Chemistry, Columbia University, New York, New York

AND

SEYMORE HOCHBERG

E. I. du Pont de Nemours and Company, Wilmington, Delaware

Received October 18, 1948

This paper represents a review of studies made during 1943, 1944, and 1945 by the Central Aerosol Laboratories at Columbia University under NDRC contract and several cooperating agencies, including the U. S. Department of Agriculture, the Army, the Navy, the British Colonial office, and others.

The studies include several steps from fundamental research to field application and illustrate the spectacular results that can be obtained by the application of fundamental principles in this field, which had relied so much on empirical information.

The precision of the work here reported and the reliability of the results were made possible largely by the development by Sinclair and La Mer (13) of a laboratory generator which produced aerosols very uniform in particle size. For these experiments modifications were made in that generator to replace the spark source of nuclei, which produced toxic oxides of nitrogen, to make the output constant with time, and to extend the range of particle diameter up to 40 microns.

Figure 1 shows the homogeneous aerosol generator as used in these laboratory experiments. The nuclei were provided by heating a nichrome wire coil in which there was a plug of sodium chloride to a temperature just below red heat. A constant-voltage transformer was required to maintain constancy in the supply of nuclei for condensation.

The primary adjustment of particle size is obtained by adjustment of the heater temperature, higher temperatures resulting in larger particles. Necessarily, therefore, with increased particle size there was also an increased mass concentration. A bimetallic thermostat controlled the temperature of the box surrounding the first heater to ±0.5°C. and the flask to a much smaller tolerance.

The temperature of the second heater is not critical but it must be higher than that of the first to prevent undue condensation and must not be high enough to result in excessive decomposition. In the case of DDT in lubricating oil, this must not exceed 230°C. The output of the generator is assessed by passing its efflux through a glass wool filter in a drying tube and either weighing the filter or determining the DDT on the filter by chemical means.

---

[1] Presented at the Symposium on Aerosols which was held under the joint auspices of the Division of Physical and Inorganic Chemistry and the Division of Colloid Chemistry at the 113th National Meeting of the American Chemical Society, Chicago, Illinois, April 22, 1948.

MORTALITY OF MOSQUITOES IN CONFINED ATMOSPHERES

In the first experiments (January, 1944) with mosquitoes (10) it was desired to test the effect of particle size on the toxicity of DDT aerosols to *Aedes aegypti* mosquitoes under conditions where there would be no relative motion of aerosol and insects. A homogeneous aerosol was confined in a wooden box of 85-l. capacity having glass windows. The aerosol was stirred first with a small fan at slow speed and then the cage of mosquitoes was introduced. A beam of light passing through the chamber onto a photronic cell indicated the decrease in aerosol concentration that might occur during the course of an exposure.

For each particle size several exposures were made at different values of time t and also of concentration C. The logarithm of the product Ct to kill 50 per cent of the female insects, $\log Ct_{50}$, is related over a considerable particle size range to the particle diameter by the expression:

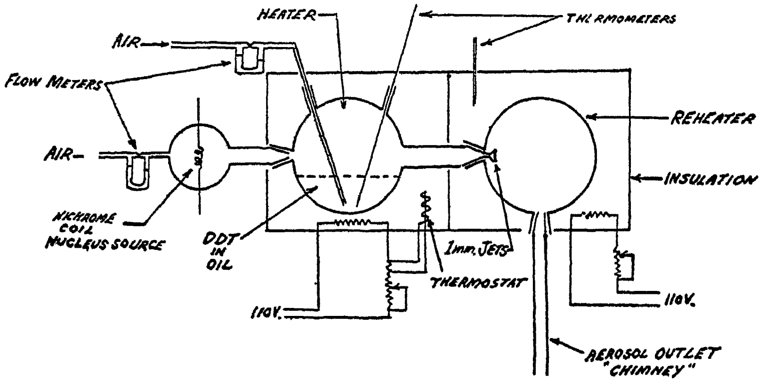


FIG. 1. Homogeneous aerosol generator

$$\log Ct_{50} = -2 \log d + \text{Constant} \tag{1}$$

Log Ct is plotted against $\log r$ [$\log r = \log d + \log 2$] in figure 2. The curve has a slope of -2, supporting the hypothesis that deposition on the insect by settling of the particles according to Stokes' law was in operation. The leveling off of the curve was believed due to the statistical effect, which would become evident at such a particle radius that a very small number of particles would be sufficient to kill an insect. A median lethal dose can be calculated then, based solely on gravity fall, assuming a cross-sectional area of 0.03 cm.^{2} for a mosquito as 3.6×10^{-6} mg. of 8 per cent DDT, or 2.9×10^{-7} mg. of DDT. This is $\frac{1}{100}$ the quantity believed on the basis of subsequent experiments to be more reliable.

The conclusions drawn from the data were that particle size should be about 10 microns diameter for optimum results. A large-scale aerosol generator was designed and built in accordance with this recommendation. The conclusion that the toxicity was proportional to the quantity depositing upon the insect was justified in spite of the fact that, as was learned soon thereafter, much of the aerosol probably reached the insects by way of convection currents and by way of their own motion rather than by gravity settling alone.

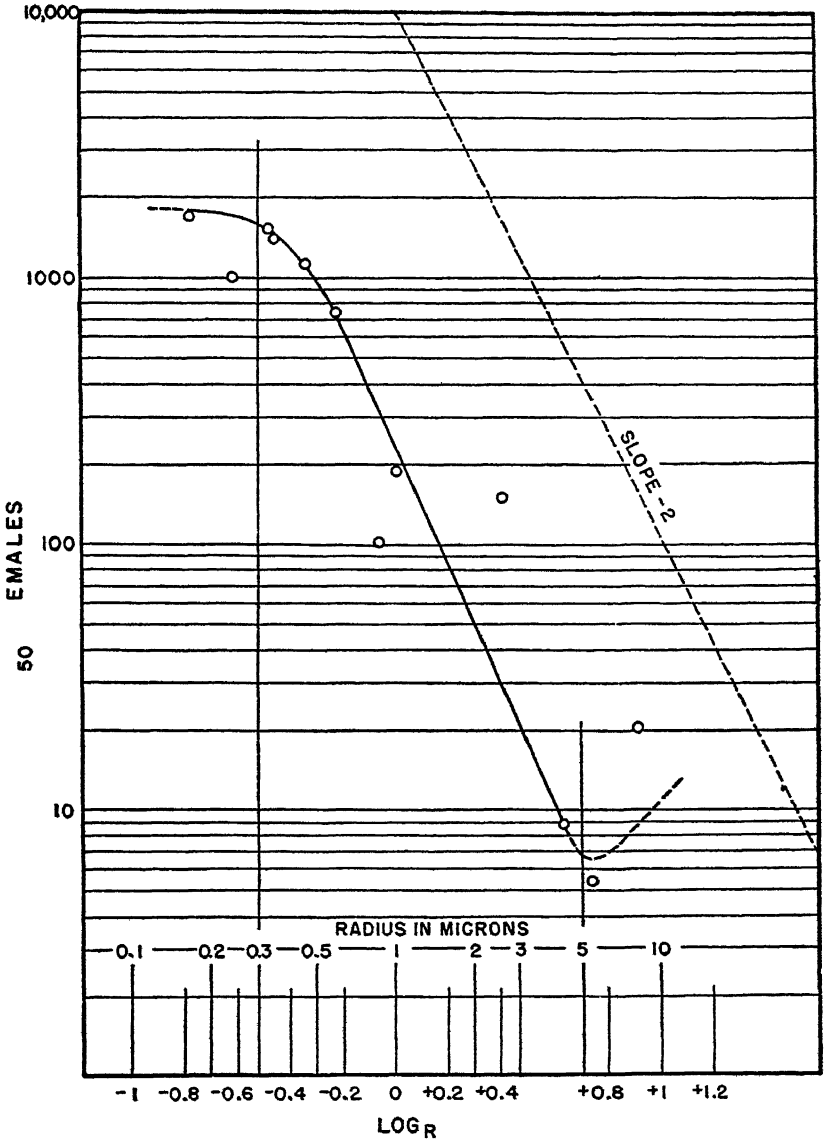


FIG. 2. Plot of log *Ct* against log *r* (La Mer *et al.*: J. Colloid Sci. **2**, 539 (1947))

THE MORTALITY OF MOSQUITOES AS A FUNCTION OF WIND VELOCITY AND PARTICLE SIZE

In order to test the effect of wind velocity and to obtain further data on the effect of particle size on the toxicity of DDT, more extensive experiments were performed in which mosquitoes in cages were exposed to aerosols within a wind tunnel 1 ft. square in cross-section (11).

Wind velocities of 2, 4, 8, and 16 miles per hour and particle sizes from 1 to 20 microns in diameter were used. For each particle size and wind velocity, enough exposures were made to determine the quantity of aerosol, *M* mg. per square foot of tunnel cross-section, required to kill 50 per cent of the female mosquitoes. This quantity was found to be determined by the product of the square of the diameter D and the wind velocity v (see figure 3).

The data obey the following law:

$$\log M = K \log D^2v + \text{Constant} \tag{2}$$

In the range $D^2v = 2$ to 200 microns² miles per hour, the data obey this relation with $K = -1$, M ranging between 300 mg. and 1 mg. per square foot. Between values of D^2v of 200 to 800 the value of K decreases to zero, and between values of 800 and 7000, K is zero, the value of M being 0.8 mg. of DDT per square foot.

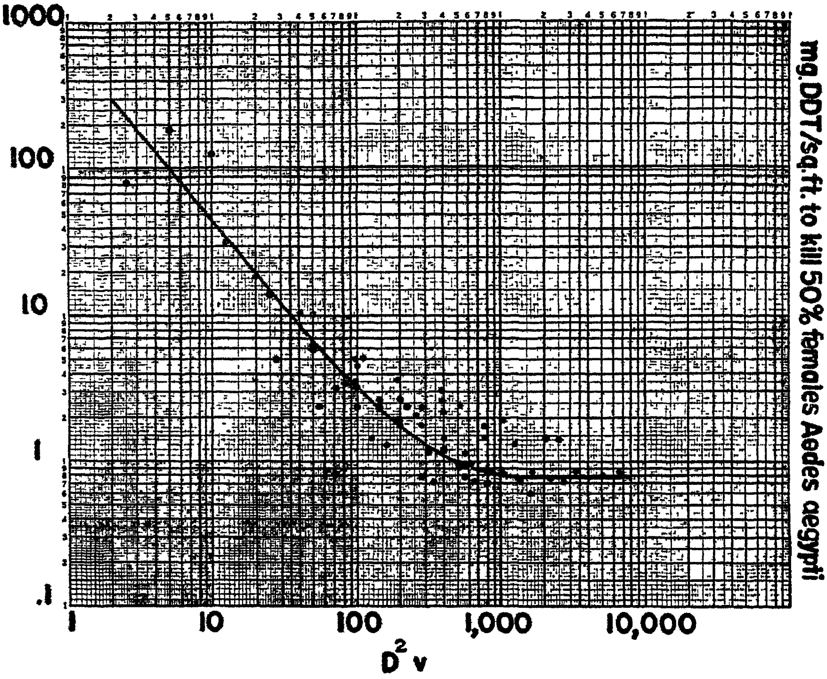


FIG. 3

That the spread of points about the curve of central tendency is due to the limited number of test insects is indicated by figure 4, where the mortality for males is plotted against that for females, each point representing a percentage mortality for a single cage containing roughly 50–100 insects. It is to be assumed that, had a sufficiently large number of insects been used, the data would have approached closer to the curve of figure 4 and also the curve of figure 3. Nevertheless, practically all the points lie within a factor of $\frac{1}{2}$ to 2 times the dose given by the curve drawn in figure 3. Use of mortality figures for males does not yield any better curve of the type of figure 3 than does the mortality curve for females, except that the former are more sensitive to DDT.

The results of these experiments were interpreted as follows: The motion of particles in an aerosol is governed principally by the motion of the air which supports them. This air, when a stationary object is in its path, tends to stream around the object and to carry the aerosol particle with it. The particle, on the other hand, being heavier than air and therefore having more inertia, tends to continue in a straight line. As a result, the path of the aerosol particle has a smaller curvature than the path of the air in which it is suspended. Therefore, when an aerosol is blown against an object a fraction of the aerosol strikes an object in its direct path, this fraction becoming larger as the velocity of the aerosol is increased, as the particle diameter is increased, or as the size of the object is reduced. When this fraction approaches unity, further increase in particle size or wind velocity ceases to have effect.

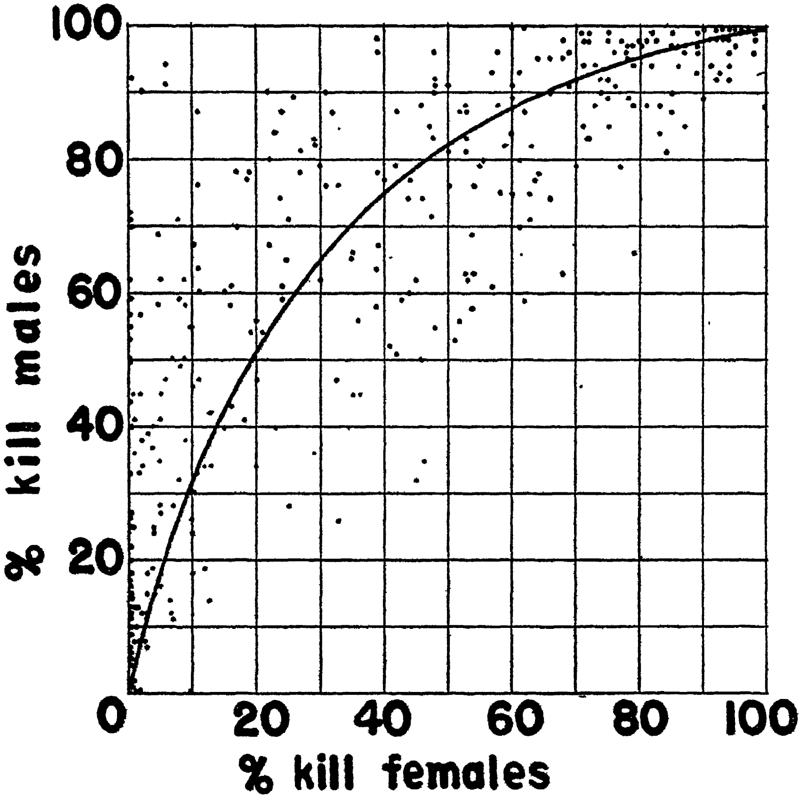


FIG. 4. Plot of mortality for males against that for females

A mathematical analysis which predicts this type of dependence of deposition upon D^2v has been given by Sell (12; see also 4 and 15) for simple geometrical objects.

The data on insects in the confined atmospheres must be interpreted as being complicated by a small velocity which resulted on account of (*a*) a small amount of flying, (*b*) putting the cages into the aerosol and taking them out, and (*c*) a certain amount of convection.

The following conclusions were drawn from the wind tunnel experiments:

1. The data indicate that the toxicity is determined by the quantity deposited on the insects which, according to Sell, would be expected in the limits of low and high values of D^2v to yield the slopes obtained in figure 3.

2. The data for larger particle size and wind speed were taken to indicate saturation, i.e., that the insects were taking up a quantity of aerosol equal to that which otherwise would have passed through their cross-section. The median lethal dose of DDT as yielded by this interpretation is 3.0×10^{-8} g. of DDT.

3. A wind tunnel could be used as a chamber for testing relative toxicity of insecticides in conjunction with an atomizer which would make particles within a certain range such that D^2v will be above 800 and yet the loss of particles by settling in the tunnel before the insect cages are reached would be negligible.

Experiments were performed in the wind tunnel using a specially made atomizer to test the toxicity of DDT in a variety of solvents (8).

It was found that DDT was most effective when it was present with a relatively non-volatile solvent, but that it lost effectiveness if used with solvents that would volatilize to relative completion before the aerosol contracted the insects. This type of toxicity determination is so convenient in operation and is so rapid that we have suggested that it be studied further as a substitute for the standard Peet-Grady chamber method.

4. In using aerosols for the kill of mosquitoes by contact with the aerosol, it is futile to use values of particle diameter such that D^2v would be greater than 700 microns^2 miles per hour. Larger values of D^2v would kill mosquitoes no better and would not have the ability to be carried into crevices and around objects as well. Under practical conditions this indicated a particle size of approximately 10 microns diameter.

GENERATION OF AEROSOLS FOR FIELD CONTROL

The laboratory data showed that it would be feasible to control insects over large areas by emission of a DDT aerosol from a generator moving crosswise to the wind and allowing the wind to carry it over long distances to the insects. However, particle size would have to be controlled carefully. A generator was developed (U. S. patent 2,416,256) that would produce particles in the most desirable particle-size range in outputs up to about 135 lb. of DDT per hour, or 50 gallons per hour of 35 per cent DDT. This machine pumped DDT dissolved in oil and water either separately or as an emulsion under pressure through heating coils where all of the water was converted into superheated steam and the mixture was ejected through a nozzle. The nozzle consisted of a tube about 2 in. long and $\frac{1}{8}$ in. in diameter. By adjustment of temperature and pressure, the average particle diameter could be adjusted from about 2 microns up to 60 microns.

The carrying power of the aerosol produced was measured by dissolving a dye in the oil, generating an aerosol containing the dye, and measuring the deposit of aerosol collected on suitable surfaces such as glass slides or plates at convenient distances downwind through various types of terrain and vegetation by colorimetric methods. Original data and plots similar to figure 5 are given in reference 1. Figure 5 gives the results of two of the typical experiments. It was found that in general the weight m per square foot deposited on a plate is an inverse exponential function of the distance x from the generator, so that

$$m = m_0 e^{-2.3kx} \qquad (3)$$

where k is the slope of $\log m$ *vs.* x and m_0 is the deposit per unit area at very short distances ($x = 0$).

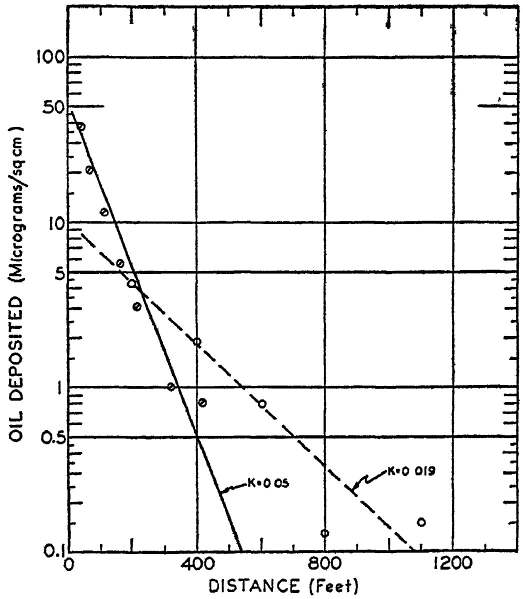


FIG. 5. Deposition as a function of distance. ○, test I (light forest); ⊘, test B (light-dense forest). Log $M = -Kd +$ Constant.

It is instructive to consider the simplified case where the ground surface is made up of surfaces similar to the plates used for the measurement and that loss occurs entirely on these surfaces.

Suppose that R is the output rate of a generator in pounds per foot of travel measured perpendicular to the direction of the wind, and m is the deposit rate in pounds per square foot.

The amount A lb. deposited in an area 1 ft. wide bounded by the generator travel line and a distance x is calculated by integrating mdx between the limits zero and x giving:

$$A = \frac{m_0}{2.3k} \quad \frac{m_0}{2.3k} e^{-2.3kx} \qquad (4)$$

If the distance x is taken as extremely large, then A becomes identical with R, the output rate per foot of travel, and

$$R = \frac{m_0}{2.3k} \qquad (5)$$

Suppose it is desired to apply a dose of m_1 at a distance x from the generator. Then from equation 3

$$m_1 = m_0e^{-2.3kx_1} \tag{6}$$

and from equation 5

$$m_1 = 2.3kRe^{-2.3x_1} \tag{7}$$

or

$$\frac{R}{m_1} = \frac{e^{2.3kx_1}}{2.3k} \tag{8}$$

Table 1 gives values for R/m_1 as a function of x_1 and k.

TABLE 1

Values of R/m_1 as a function of x_1 and k

x_1	R/m_1	
	$k=0.0001$	$k=0.005$
0	434	87
100	546	276
200	687	870
500	1370	27,600
1000	4340	8,700,000

TABLE 2

Most economical value of k to produce a given deposition rate at a distance x_1 together with required dosage

x_1	k	R/m
0		0
100	0.00434	272
200	0.00217	545
500	0.00087	1360
1000	0.000434	2720

Differentiation of equation 8 with respect to k yields an equation for the most advantageous value of k, k_{econ}, to yield lowest values of R for a given value of m at a distance x_1.

$$k_{\text{econ}} = 1/(2.3x_1); \qquad (R/M)_{\text{econ}} = e/2.3k_{\text{econ}} \tag{9}$$

Table 2 reveals that the output becomes proportional to the distance desired provided k is always optimum.

Comparison of output rates R between tables 1 and 2 shows that the value of k has a tremendous effect upon the amount of material required to kill insects at a long distance. Only through proper particle-size control can output values be kept sufficiently low if long-distance kill (over a few hundred feet) is desired.

FIELD TESTS

Field tests were performed under various conditions of terrain from open country to jungle and forested areas under various weather conditions to test these hypotheses (1, 2, 3, 5, 6, 7, 9, 14).

Selected tests as described below will indicate the power of controlled particle size.

I. The generator was transported for 500 ft. along the edge of a fringe of dense vegetation bordering a light forested area crosswise to the wind. Wind speed in the forest was 0.3 mile per hour or less. The dosage was estimated as 15 gallons of emulsion of 5 per cent DDT (total 5 lb.) for the 500-ft. front. Adult mosquitoes were trapped in cages and in barrels and were placed at seven positions from 100 to 1100 ft. at ground level and also at 6 ft. above ground level.

Pans of larvae were also placed at these positions. Table 3 gives the mortality figures obtained. A peculiar topographical condition was blamed for the relatively low deposition and poor kill at 800 ft. The dosages given in the table are those collected on a horizontal glass plate. The actual consumption of DDT can be assumed to be much greater, being absorbed on all types of surfaces (tree leaves, etc.), of which there can be many square feet per square foot of ground area. k in this case was 0.019 (see figure 5, test I).

TABLE 3

Forested area

DISTANCE	DDT	PER CENT KILL LARVAE	PER CENT KILL ADULTS	
			Barrels	Cages
feet	*lb./acre*			
100		100	100	100
200	0.038	100	100	100
400	0.017	100	100	100
600	0.0069	100	100	100
800	0.0011	100	190	60
1100	0.0013	95	100	85
1200				100

II. Another test was made in rather open marshy country covered with a low weed growth. The wind speed was 2 miles per hour, the generator path was 500 ft. long, and it emitted 0.5 lb. of DDT per 1000 ft. of front as an aerosol of 3–4 microns diameter. Counts were made of landing rates on a man's back over several previous mornings of 200–300 mosquitoes per minute. Stations were laid out at 1400, 1700, and 2400 ft. downwind from the generator. Complete elimination of free mosquitoes was indicated at the end of the run and all caged mosquitoes were dead or suffering visibly from poisoning 3 hr. after the test (see table 4). In addition, landing rates were recorded up to 5300 ft. Comparison of the figures with checks made upwind from the generator indicates that *100 per cent mortality was obtained for half a mile and at least 95 per cent mortality was obtained for an additional half mile.*

III. A set of two tests at Mantoloking, New Jersey, a resort town on the Atlantic shore, illustrates again the importance of particle size. Mantoloking is part of a sand spit extending north and south on the New Jersey coast and is separated from the mainland by Barnegat Bay. At Mantoloking the spit is 800 to 2000 ft. wide. In the first test, an aerosol of 10 microns diameter was used, emitted from the generator mounted on a truck. With an output rate of 34 lb. of DDT per mile, there was a complete elimination of adult mosquitoes in the area, but larvae in jars were hardly affected even at a distance of only 50 ft. In the second test, the generator was mounted on a skiff operating in the bay.

TABLE 4

Marshy country

DISTANCE	LANDING COUNTS		CAGED MOSQUITOES, PER CENT MORTALITY AFTER 3 HR.*
	Before test	3 hr. after test	
feet			
1400	35	0	90
1700	50	0	95
2400	112	0	90

* All living mosquitoes affected by DDT.

TABLE 5

Test B: Mantoloking

STATION NO.	34 LB. DDT PER MILE POSITION	PER CENT MORTALITY	
		Adults	Large larvae
1	Bay	70	100
2	Middle	100	100
3	Ocean	80	100
4	Bay	100	100
5	Middle	100	100
6	Ocean	100	100
7	Bay	100	96
8	Ocean	64	76

A steady wind blowing from the bay served to spread the aerosol over the land. Caged adults and larvae in jars were placed at eight positions, three on the ocean beach, on the far side of the spit, three at the shore of the bay nearest the path of the generator, and two others part of the way across the spit. Wind velocity was 3–5 miles per hour. Particle size for this test was 10–25 microns diameter. Table 5 shows the mortality figures.

Output was 34 lb. of DDT per mile, or the same as in the previous test. Larvae were killed this time even at the extreme distance, approximately 2000 ft. It is felt that doubling the quantity emitted without change of particle size would have provided 100 per cent control.

IV. Another experiment was conducted (3, 14) jointly with C. B. Symes and

L. D. Hadaway, the entomologists of the British Colonial Office in Georgetown, British Guiana. The aerosol generator was operated for a path of 1000 ft. on each of three streets which composed Lodge Village, a malaria-infested community on the outskirts of Georgetown. The area south of the village consisted first of 2000 ft. of cemetery, then 2000 ft. of rice fields, and then 4000 ft. of cane fields. Vegetation grew thickly to a height of 20 ft. around the houses. Rice fields were covered with vegetation about a foot high. The sugar cane was 6–8 ft. high. Sixty stations were laid out in this area up to 8000 ft. downwind from the generator. Extensive daily counts of 500 to 1400 mosquitoes per room were made in the native houses of the village. Complete kill of caged adults and almost complete kill of free adults up to 2300 ft. beyond the village was evident after the application of a total of 36.6 lb. of DDT. Up to 4000 ft. the kill diminished to 50 per cent and up to 8000 ft. went to zero. *Larvae were killed at all stations up to 8000 ft.* except in one most easterly station.

SUMMARY AND CONCLUSIONS

1. Data on the effect of particle diameter and wind velocity show that the toxicity of DDT aerosols to mosquitoes follows the law of deposition of aerosols on geometrical objects, as calculated by Sell.
2. A method of determination of toxicity is proposed making use of a wind tunnel instead of a Peet-Grady chamber. Problems of contamination can be largely avoided and rational results in terms of absolute toxicity obtained.
3. The deposition from an aerosol is an inverse exponential function of the distance from the source. The constant involved depends upon a number of factors, such as turbulence, particle size, density of vegetation, wind velocity, etc.
4. This exponential relationship makes necessary careful control of particle size as well as selection of favorable meteorological conditions if good results are to be achieved.
5. A generator was developed by which large-scale control of mosquitoes by ground dispersal is possible over difficult and inaccessible terrain. This was achieved through selection of particle size and by operation under favorable meteorological conditions.

REFERENCES

(1) BRESCIA *et al.*: J. Econ. Entomology **39**, 698 (1947).
LA MER, BRESCIA, WILSON, HODGES, AND ROWELL: OSRD Report No. 573 (September 18, 1945); J. Econ. Entomology **40**, 309, 313, 316 (October, 1947).
(2) BRESCIA AND WILSON: "Reports on the Use in the South Pacific of the Navy Screening Smoke Generator Converted to an Insecticidal Aerosol Generator according to the Hochberg–La Mer Principle," OSRD Report No. 5730. Also J. Econ. Entomology **40**, 309, 313, 316 (October, 1947).
(3) HODGES AND ROWELL: "Mosquito Control with the Hochberg–La Mer Aerosol Generator in British Guinea," OSRD Report No. 5310.
(4) JOHNSTONE, WINSCHE, AND SMITH: Chem. Revs. **44**, 353 (1949).
(5) LA MER: "Efficacy of DDT and DNOC as Insecticides for Grasshoppers, etc.," OSRD Informal Report 10.2–25 (October 25, 1945).
(6) LA MER: "Black Fly Control with DDT-Oil Aerosols at Lake Placid, New York," OSRD Informal Report 10.2–26 (November 8, 1945).

(7) La Mer, Hochberg, *et al.*: "Field Tests against Salt Marsh Mosquitoes at Mantoloking, New Jersey," OSRD Informal Report No. 10.2-23 (October 23, 1945).
(8) La Mer *et al.*: OSRD Report No. 5936.
(9) La Mer *et al.*: "Cankerworms as Test Insects for DDT Field Tests of Models of the Hochberg-La Mer Aerosol Generator at Hempstead Lake State Park," OSRD Report No. 6004 (September 28, 1945).
(10) La Mer *et al.*: J. Colloid Sci. **2,** 539 (1947).
(11) Latta, La Mer, *et al.*: "Effect of Particle Size and Speed of Motion of DDT Aerosols of Uniform Particle Size in a Wind Tunnel on the Mortality of Mosquitoes," OSRD Report No. 5566 (July 30, 1945); J. Wash. Acad. Sci. **37,** 397 (1947).
(12) Sell: *Forschungsarbeiten Verein deutscher Ingenieure*, Vol. 1, p. 347. V.D.I. Verlag, Berlin (1931).
(13) Sinclair and La Mer: Chem. Revs. 44, 245 (1949).
(14) Symes and Hadaway: Reports to the British Colonial Office.
(15) Winsche: OSRD Informal Report No. 10.4-55 (March 1, 1944).

THE DISPERSION AND DEPOSITION OF AEROSOLS[1]

H. F. JOHNSTONE, W. E. WINSCHE,[2] AND L. W. SMITH[2]

Noyes Chemical Laboratory, University of Illinois, Urbana, Illinois

Received October 18, 1948

The dispersion and deposition of aerosols is considered on the basis of Sutton's theory for the diffusion of gases in the lower atmosphere and Sell's work on the efficiency of deposition of particles by impaction. An estimate is made of the rate of deposition of particles dispersed from a line source under various meteorological conditions. Application of the impaction curves has been made to the penetration and deposition of insecticidal aerosols through a forest canopy. Under conditions of low wind and in the absence of a downdraft, the theory predicts that the larger particles will penetrate the canopy better than the small. The opposite is true, however, when the dissemination depends on the downdraft from a plane. Sell's curves may also be used to predict the effect of the drop size of aerosols on the mortality rate of moving and stationary insects. The calculated results agree closely with the experimental observations of La Mer and coworkers.

The dissemination of particles suspended in the atmosphere is a problem of great practical importance. The dispersal of smokes and fumes from industrial plants is of concern to public health authorities and has received much attention (2, 8). The mathematical theory of atmospheric diffusion and its application to military problems in chemical warfare have been discussed in a recent paper by Sutton (8). Insecticides are often dispersed in the form of aerosols of small solid or liquid particles. For maximum efficiency these must be deposited uniformly on the ground or water surfaces, or on vegetation inhabited by the insects to be destroyed. Occasionally it is desirable to combat insects while they themselves are flying. Direct contact with aerosol particles by impaction on the wing or body then becomes important. Optimum results are obtained by dispersing the aerosol in particles containing approximately the lethal amount of the insecticide. An examination of some of the fundamental concepts of aerosols has proven valuable when interpreting results obtained in the laboratory and field, and in predicting the effects to be obtained under various meteorological and topographical conditions.

DEPOSITION IN OPEN AREAS

When an aerosol is dispersed from a generator the particles are carried by the wind with little tendency to settle. It can be shown readily that the rate of fall of particles, even as large as 50 microns diameter, is small compared to the normal turbulent diffusion processes of the atmosphere. Deposition on the ground takes place therefore, not by settling through the height of the cloud, as has frequently been supposed, but by the particles actually being brought to the ground, just as the molecules of a gas cloud are mixed and brought to the ground by eddy diffusion. Adjacent to the ground is a more or less stagnant layer through which aerosol particles must then pass by true settling, or by Brownian diffusion. The latter may be neglected for particles of diameter larger than 0.5 micron. Thus not all of the particles that come close to the ground are deposited. Some are carried aloft again by wind currents. The rate of deposition of the aerosol particles per unit area is obviously the product of the concentration at ground level and the settling velocity.

---

[1] Presented at the Symposium on Aerosols which was held under the joint auspices of the Division of Physical and Inorganic Chemistry and the Division of Colloid Chemistry at the 113th National Meeting of the American Chemical Society, Chicago, Illinois, April 22, 1948.

This paper is based in part on work done for the Office of Scientific Research and Development under Contract OEMsr-102 with the University of Illinois.

[2] Present address: Brookhaven National Laboratory, Upton, Long Island, New York.

The problem of diffusion of gases in the lower atmosphere has been treated extensively in the recent papers by Sutton (7, 8). While the theoretical expressions cannot be claimed to give highly accurate values for concentrations of gases released into the air because of the widely varying conditions in the winds, they are reasonable guides when observations on the wind gradient are made above the layer in which the effects of surface roughness predominate. In order to use Sutton's equations for aerosols which settle out, a correction must be made for the amount of material which is deposited. For particles larger than 1 micron this correction becomes important. The following derivation applies to a cross wind line source but analogous considerations may be applied to a continuous point source.

The equation for the downwind concentration of a gas at a distance x from a line source of strength W is[3]

$$C = \frac{W e^{-z^2/c_z^2 x^m}}{\sqrt{\pi}\, c_z\, u x^{m/2}} \tag{1}$$

---

[3] Nomenclature: consistent units are used except as indicated in the text.

c_z = vertical diffusion coefficient depending on meteorological conditions
C = concentration of cloud
Ct = concentration–time product; subscript g refers to gas cloud, a to aerosol cloud
D = dimension of impacting object, diameter of sphere
d = diameter of particle
f = fraction of particles remaining airborne
g = acceleration of gravity
h = height of source above ground
k = proportionality factor relating drag resistance to velocity of particle
M = lethal amount of insecticide
m = a meteorological constant
n = number of units of distance; also number of particles of insecticide containing lethal dose
Q = number of particles crossing vertical plane of unit area
t = time
u = mean wind velocity
v = velocity of downdraft
v_s = rate of settling of particle
W = line source strength, mass of agent emitted per unit time per unit distance
W_{n^+} = probability that n or more particles will strike a given area in time t
w = mass of particle
x = horizontal distance

The quantity c_z is a generalized diffusion coefficient which depends on meteorological conditions, the gustiness of the atmosphere, and only slightly on u, the mean wind speed. Gustiness is obviously a variable quantity and depends upon many things, including the net radiation received by the ground, the topography, and the height of observation. The exponent m is also a meteorological constant which may be measured by the wind profile. Its value varies from approximately 1.85, for conditions of high turbulence due to a large lapse rate, to about 1.4 for inversion conditions such as frequently occur in the late evenings or at dawn. Equation 1 does not take into account the effect of the ground on the gas concentration. Allowing for the reflection of the cloud at the plane $z = 0$, when the source is at the height h, the equation becomes

$$C = \frac{W}{\sqrt{\pi}\, c_z u x^{m/2}} [e^{-(z-h)^2/c_z^2 x^m} + e^{-(z+h)^2/c_z^2 x^m}] \tag{2}$$

Now in an aerosol cloud in which deposition occurs by settling across the stagnant layer adjacent to the ground, reflection is not complete. Furthermore, if the particles are large there is some tendency for settling to take place even in the turbulent region. Introducing these corrections the equation becomes

$$C = \frac{W}{\sqrt{\pi}\, c_z u x^{m/2}} [e^{-\left(z+\frac{v_s x}{u}-h\right)^2/c_z^2 x^m} + \alpha e^{-\left(z+\frac{v_s x}{u}+h\right)^2/c_z^2 x^m}] \tag{3}$$

Here α is the average reflection coefficient up to the distance x; that is, it is that fraction of the material in the cloud brought to the ground that is not deposited.

The average value of α may be determined by the material balance:

$$\frac{1}{W}\int_0^x C_0 v_s \, dx = 1 - \frac{1}{W}\int_0^\infty Cu \, dz \tag{4}$$

Equation 4 states that the fraction deposited is the fraction that does not diffuse across an infinite vertical plane at the distance x from the source. If the line source is at ground level, and if the effect of settling in the turbulent region can be neglected, equation 4 gives

$$\frac{(1+\alpha)v_s}{\sqrt{\pi}\, c_z u}\int_0^x \frac{dx}{x^{m/2}} = 1 - \frac{(1+\alpha)}{\sqrt{\pi}\, c_z x^{m/2}}\int_0^\infty e^{-z^2/x^m c_z^2}\, dx \tag{5}$$

z = vertical distance
α = reflection coefficient fraction of aerosol cloud brought to the ground by turbulent diffusion that is not deposited
γ = characterization constants for foliage
ϵ = efficiency of deposition of aerosol particles on objects by impaction
μ = viscosity of air
φ = average number of particles striking given surface per unit time
σ = surface area of insect for deposition of aerosol particles; subscript h refers to resting insect; f to flying insect; v to frontal impact area of flying insect
ρ = density of aerosol particles

or

$$\frac{(1+\alpha)v_s}{\sqrt{\pi}\,c_z u}\,\frac{x^{1-m/2}}{\left(1-\frac{m}{2}\right)} = 1 - \frac{1+\alpha}{2} \tag{6}$$

Either side of equation 6 expresses the fraction deposited as a function of distance. Solving for the value of the reflection coefficient,

$$\alpha = \frac{2}{1+\dfrac{2v_s\,x^{1-m/2}}{\sqrt{\pi}\,u\left(1-\frac{m}{2}\right)c_z}} - 1 \tag{7}$$

TABLE 1

Fraction of homogeneous aerosol clouds deposited at various distances from a continuous line source on the ground

DISTANCE	WIND VELOCITY = 2 MILES PER HOUR					WIND VELOCITY = 5 MILES PER HOUR				
	Particle diameter in microns					Particle diameter in microns				
	1	5	10	20	50	1	5	10	20	50
yards										
10	0.005	0.129	0.37	0.70	0.90	0.000	0.06	0.18	0.49	0.85
50	0.007	0.154	0.42	0.74	0.93	0.002	0.07	0.22	0.54	0.88
100	0.009	0.163	0.44	0.76	0.95	0.002	0.07	0.24	0.57	0.88
500	0.010	0.195	0.49	0.79	0.96	0.004	0.09	0.28	0.61	0.90
1000	0.012	0.208	0.51	0.81	0.96	0.004	0.10	0.30	0.63	0.91
5000	0.014	0.245	0.56	0.84	0.97	0.005	0.11	0.34	0.68	0.93
10000	0.015	0.263	0.58	0.85	0.97	0.006	0.12	0.36	0.70	0.93

In order to compare the amount of deposition from aerosol clouds dispersed over an open terrain, solution of equation 6 has been made using values of $c_z = 0.12$ and $m = 1.75$ which are given by Sutton (8) for "average" conditions, corresponding to a small vertical temperature gradient and a low wind velocity. The results of the calculations are shown in table 1. It is interesting to note that, under the conditions chosen, most of the deposition takes place near the source. This is because the ground concentration in this area is very high. Aerosols with particles less than 10 microns in diameter, however, are carried for considerable distances without serious depletion of the concentration by deposition. The normal dilution of the cloud by vertical expansion decreases the ground concentration so much that little deposition occurs beyond 100 yards from the source. It should be noted that, if the source were located slightly above the ground and for smaller values of m and c_z corresponding to inversion conditions, more nearly uniform deposition can be obtained for short distances from the generator before the cloud is depleted.

HORIZONTAL PENETRATION OF AEROSOLS THROUGH FOLIAGE

It is frequently of interest to estimate how far an aerosol will penetrate through foliage, such as a forested area or low vegetation. Of course, the problem cannot be solved completely without knowing the factors that characterize the size and shape of the obstacles in the path of the aerosol. Certain generalities can be reached, however, by characterizing the denseness of the foliage in terms of vertical and horizontal lengths. It is necessary to consider the efficiency of impaction of small particles on the surfaces, as it is well known that aerosol particles tend to follow the streamlines of the gas and thus to be carried around the surfaces of obstacles placed in their path. Some interesting conclusions concerning the efficiency of various shapes of bodies for collecting dust particles from gas streams have been reached by Sell (6), from a simple analysis of the motion of dust particles. From experimental measurements on the streamlines about bodies he determined the efficiency of deposition of aerosol particles, ϵ, as a function of the group wu/kD. For small spherical particles in the Stokes' law range

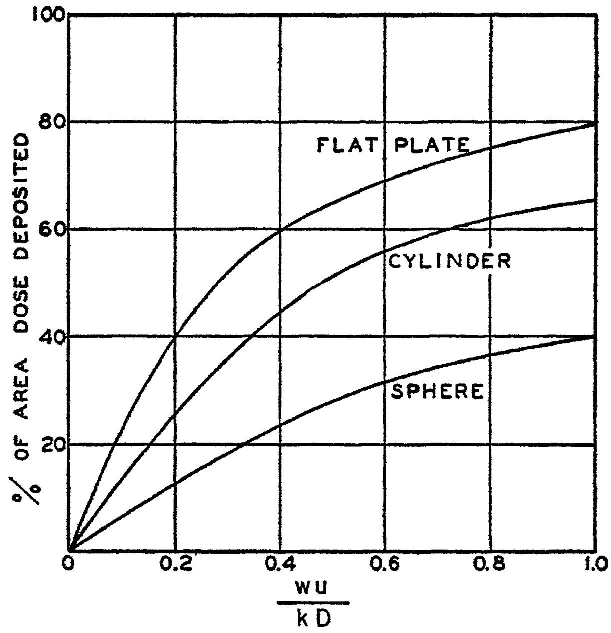


FIG. 1. Deposition of aerosols on surfaces (6)

$$k = 3\pi\mu d \tag{8}$$

and

$$\frac{wu}{kD} = \frac{1}{18}\frac{d^2 u\rho}{\mu D} \tag{9}$$

The efficiency of deposition of aerosol particles on circular cylinders, discs, and spheres is shown in figure 1.

The density of foliage may be characterized by two lengths, δ and γ, where δ is the horizontal distance for which the sum of the vertical foliage surfaces in any cross section is equal to the cross section, and γ is a similar distance in the vertical direction. The ratio δ/γ will be taken, for the purpose of illustration, as equal to 2; that is, the foliage will be considered as twice as dense in the vertical direction as in a horizontal direction. This seems to be reasonable in view of the fact that a larger proportion of the leaf surfaces are horizontal. When an aerosol penetrates horizontally through the foliage, the particles will be deposited in two ways, by impingement on vertical surfaces and by settling on horizontal surfaces. The fraction of the total number of particles lost in traveling the horizontal distance Δx is given by

$$-\frac{\Delta Q}{Q} = \frac{\epsilon \Delta x}{\delta} + \frac{\Delta z}{\gamma} \tag{10}$$

where Δz is the distance fallen in a vertical direction, and ϵ is the fraction of the area dose deposited on a vertical surface as given by Sell. We shall assume that ϵ lies between the value for a flat plate and for a circular cylinder. Of course, D will vary with the type of foliage but for purposes of illustration let us assume that it is 1 cm. If v_s is the settling velocity of the drops, we may write the above equation as

$$-\frac{\Delta Q}{Q} = \frac{\epsilon \Delta x}{\delta} + \frac{v_s}{u}\frac{\Delta z}{\gamma} \tag{11}$$

If we let n be the number of units of δ traveled downwind, that is, the number of times that a small shot on the average would penetrate through a surface in traveling the distance x, $n = x/\delta$, we have

$$-\frac{dQ}{Q} = \epsilon \, dn + \frac{v_s}{u}\frac{\delta}{\gamma} \, dn \tag{12}$$

The solution of this is $Q = fQ_0$, where f is the fraction penetrating to a distance n downwind, and

$$f = e^{-(\epsilon + (v_s/u)(\delta/\gamma))n} \tag{13}$$

The quantity f has been evaluated as a function of n for various size drops and for wind speeds of 5 and 1 miles per hour. The results are shown in figure 2. In using units of n, no particular denseness of the foliage has been assumed. Reasonable values of δ for dense and light foliage are perhaps 5 and 100 ft., respectively.

VERTICAL PENETRATION

A similar consideration of the vertical penetration of an aerosol through foliage leads to some interesting conclusions concerning the dispersion of aerosols over a forested area by means of an airplane, especially when the effect of the downdraft from the plane is taken into consideration. Using the same notation as in the discussion on horizontal penetration, if the aerosol is settling downward without any vertical component of the wind,

$$-\frac{dQ}{Q} = \frac{\epsilon u}{v_s} \, dn + 2 \, dn \tag{14}$$

$$f = e^{-(2+\epsilon u/v_s)n} \tag{15}$$

The ratio δ/γ is again assumed to be 2 and n is the number of units of the horizontal distance δ measured from the top of the canopy. The fraction f of the aerosol penetrating to a depth n is shown in figure 3 for drops of various sizes, and for a crosswind through the foliage of 1 mile per hour. If there is no crosswind, drops of all sizes will be deposited on the foliage to the same extent. This is shown by the broken line in the figure. It thus appears that *when there is a crosswind and no downdraft, the larger drops will penetrate downward through a forested area more efficiently than the smaller drops.* This apparent anomaly

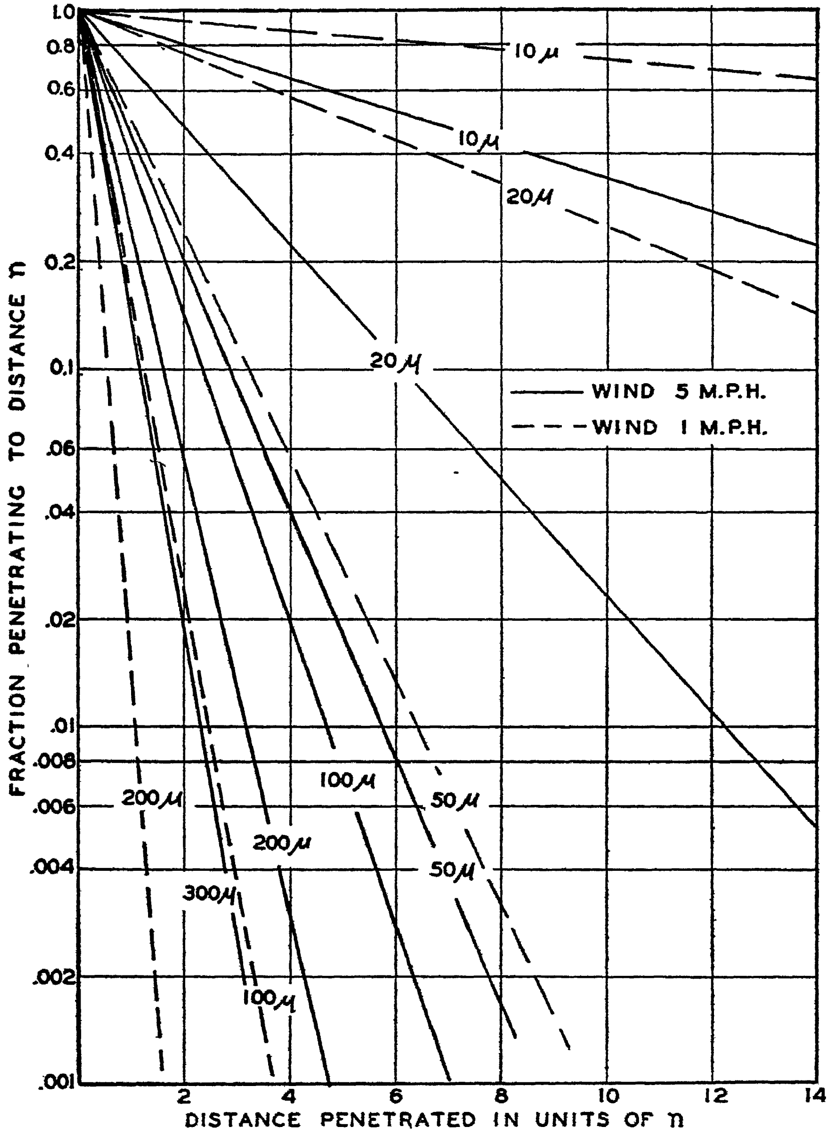


Fig. 2. Horizontal penetration of an aerosol through a forest

can be explained by examining the effect of the diameter of the particle on the factors in the exponent of equation 15. The diameter of a particle affects both the efficiency of deposition (ϵ) and the rate of fall (v_s). For diameters less than 80 microns, both numerator and denominator are approximately proportional to d^2. For larger particles, both quantities increase more slowly than d^2, the efficiency leveling off at a smaller value of d for any value of D than does the settling rate v_s, which continues to increase with the diameter. This tends to

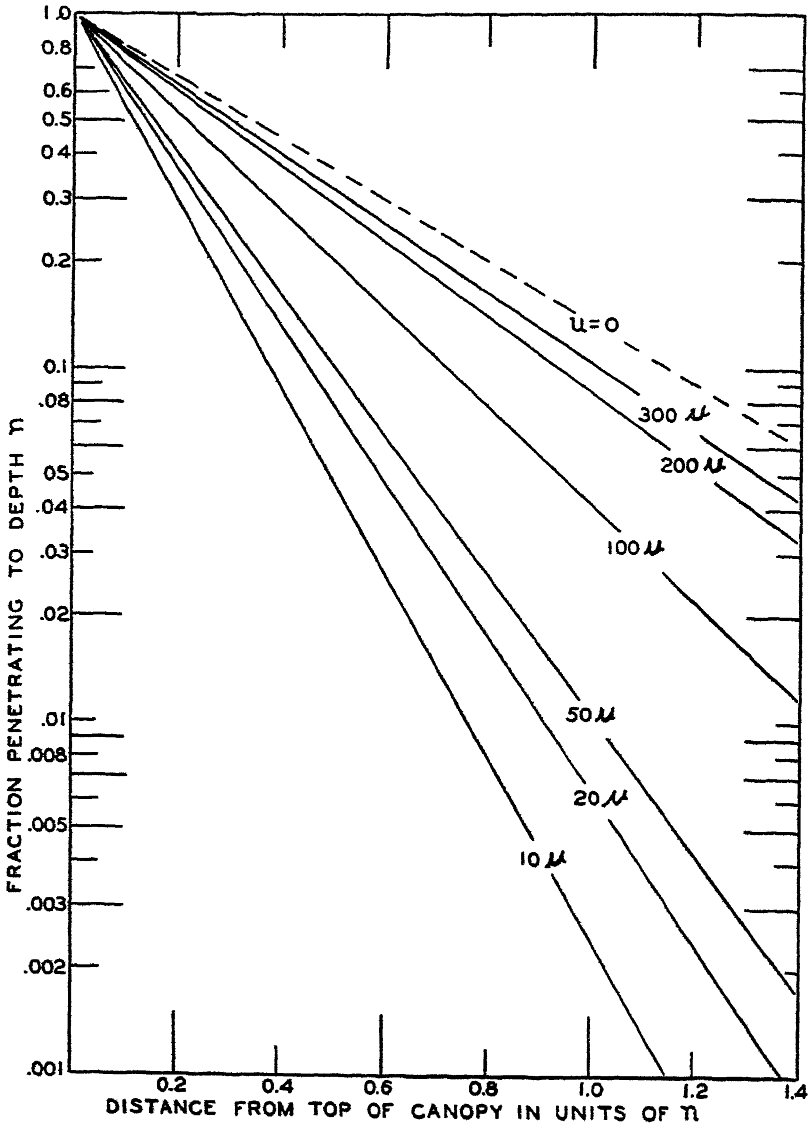


FIG. 3. Vertical penetration of an aerosol through foliage; crosswind velocity 1 mile per hour; no downdraft.

make the value of f larger for the larger drops. Physically, the effect is essentially that of the larger drops falling faster and not having as much chance to impinge on the foliage in a horizontal wind as do the smaller drops, since they are not carried past as many vertical surfaces.

If the aerosol is dispersed by an airplane which is flying only a short distance above the foliage, the downwash of the plane will drive the aerosol downward through the layers of foliage. Since the impaction on the foliage is a function of the square of the drop size, this downward component will increase the relative penetration efficiency of the smaller drops. If there is a downwash from the plane it is reasonable to assume that it dies out fairly rapidly in the foliage and that the time rate of decrease is proportional to the square of the velocity, i.e.,

$$\frac{dv}{dt} = kv^2 \tag{16}$$

where v is the downwash velocity and k is a constant of proportionality. This leads to

$$v = v_0 e^{-kz} \tag{17}$$

where v_0 = the downwash velocity at the top of the foliage canopy and v is the downwash velocity at a distance z below the top. k has the dimensions of reciprocal distance. We shall arbitrarily take $k = 1/\gamma$, so that we have

$$v = v_0 e^{-2n} \tag{18}$$

where n is again the number of units of δ below the top of the canopy.

The fraction of aerosol lost in penetrating a distance dy is

$$-\frac{dQ}{Q} = \epsilon(u)\frac{dx}{\delta} + [f\{\epsilon(v), V_s\}]\frac{dz}{\gamma} \tag{19}$$

The quantity in brackets refers to the combined fractions deposited due to settling and impingement. This quantity is not known and no attempt is made to evaluate it. We shall make the assumption that, when $v > V_s$, deposition takes place only by impingement, and, when $v < V_s$, deposition takes place only by settling. This approximation is admittedly rough, but it will give an indication of the true state of affairs. Also, since the form of $\epsilon(u)$ is not known and v contains y explicitly, we cannot integrate the above equation directly, so we define an average velocity $\bar{v}$ where

$$\frac{\int_{n_1}^{n_2} v_0 \epsilon^{-2n}\, dn}{n_2 - n_1} \tag{20}$$

or

$$\bar{v} = \frac{v_0}{2(n_2 - n_1)}\left(e^{-2n_1} - e^{-2n_2}\right) \tag{21}$$

and calculate the penetration step by step, using

$$f_i = f_{i-1}\, e^{-\left(\epsilon(\bar{v})\frac{\delta}{\gamma} + \epsilon(u)\frac{u}{\bar{v}}\right)^{(n_i - n_{i-1})}} \qquad \text{for } \bar{v} > V_s \tag{22}$$

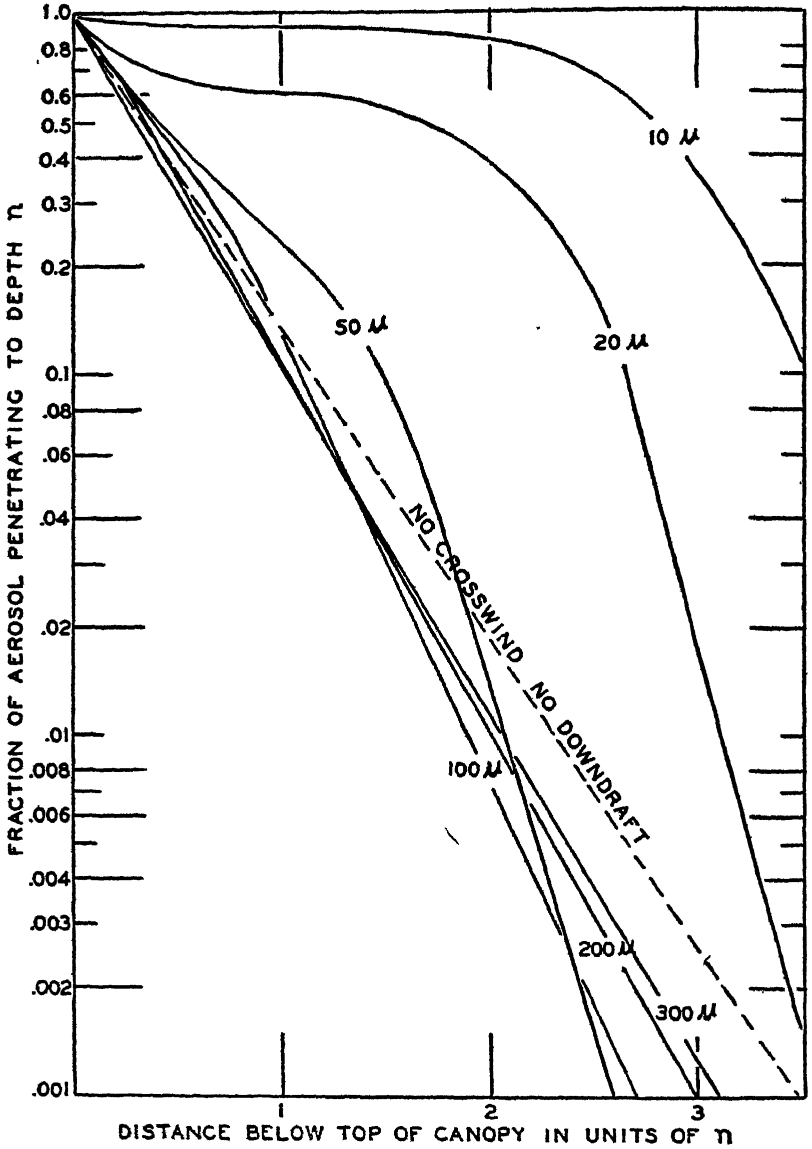


FIG. 4. Vertical penetration of an aerosol through foliage in the presence of a downdraft. u, wind velocity, 1 mile per hour; v_0, downdraft, 5 miles per hour.

and

$$f = f_k e^{-\left(\frac{\delta}{\gamma} + \epsilon(u)\frac{u}{\bar{v}_s}\right)(n - n_k)} \qquad \text{for } \bar{v} < V_s \qquad (23)$$

where f_k is the value of f at the point where $\bar{v} = V_s$.

The ratio f has been evaluated as a function of n, using increments of $n = 0.2$ for calculation. The results for various drop sizes are shown in figure 4. The solid lines are for a crosswind of 1 mile per hour and a downdraft of $v_0 = 5$

miles per hour. The dashed line shows the penetration for $v_0 = 0$ and $u = 0$. The effect of the downwash in increasing the penetration efficiency of the smaller drops is quite evident. After a certain depth has been reached, the downdraft disappears and the drops deposit only by settling. Here the larger drops become more efficient in penetration.

The optimum position of a spray outlet on an airplane to get maximum penetration into foliage is the point where the downdraft is utilized to best advantage and the smallest percentage of the droplets are carried into the turbulent wake. This position should be forward of and below the trailing edge of the wing where the flow lines have maximum divergence. If the spray is too close to the trailing edge or to the lower surface of the airfoil, it will enter the wake and be dissipated in turbulent motion. If the spray is too far below the airfoil, the downward component of velocity may be quite small.

In the discussion above, no consideration has been given to inhomogeneities, or striations, which are always present in any natural foliage. Also, any updrafts which may be present due to meteorological turbulence are neglected. The vertical components of turbulence occurring in the lower layer of the atmosphere will keep some of the smaller drops from reaching ground level. This may be an advantage in forested regions by giving a more uniform deposit on the foliage, thus leading to better control by the residual effect of the insecticide.

OPTIMUM DROP SIZE FOR CONTACT EFFECT OF INSECTICIDES

The effect of the particle size of the aerosol on the probability of hitting the insects with sufficient material to cause mortality is of interest. It is apparent that an optimum particle size must exist for each percentage kill. The application of statistical analysis to this problem was first suggested by Rodebush (5). The following treatment differs somewhat from the original and leads to some more general conclusions.

Consider the particles of an aerosol cloud falling upon a flat horizontal surface through still air. The velocity of fall depends upon the drop diameter. Consequently, different amounts will be deposited on any surface exposed to the same concentration–time product Ct_p, depending upon the diameter of the drops composing the aerosol cloud. Although each small area receives, on the average, a certain number of drops, there is a certain probability that it will receive more or less than this number. For instance, if the probability is 1/2 that of the number of drops falling on each insect is equal to or more than the number of drops necessary to cause mortality, then 50 per cent of the insects so exposed will be killed.

Three general cases will be considered:

I. The mosquito[4] is stationary and the aerosol is deposited only by Stokes' law of settling of the drops.

II. The mosquito is flying horizontally at a speed of 3 miles per hour, and in addition to the drops that settle on it, some are picked up by impingement. Impingement on the wings, which is very likely an important factor due to their rapid motion, is neglected.

---

[4] The discussion here is confined to mosquitoes but should be applicable to other insects as well.

III. The mosquito is resting on a screen through which the aerosol is passing with the wind at different velocities.

In the numerical calculations, the lethal amount of the insecticide must be known. For DDT, the lethal value for mosquitoes is between 10^{-8} and 10^{-10} g.

Case I. Resting mosquito

The average mass of aerosol deposited in unit time on a horizontal surface of area σ_h is given by

$$Cu\sigma_h = C\,\frac{d^2\sigma_h g}{18\mu} \tag{24}$$

The average number of drops deposited on σ_h per unit time is

$$\varphi = \frac{C}{3}\,\frac{\sigma_h}{\pi}\,\frac{g}{\mu d} \tag{25}$$

The cloud concentration necessary to give this number is

$$C = \frac{3\pi\mu d\varphi}{g\sigma_h} \tag{26}$$

Then the value of Ct_p necessary to have an average of φt drops strike the area σ_h in time t is

$$Ct_p = \frac{3\pi\mu d}{g\sigma_h}\cdot\varphi t \tag{27}$$

For any given value of M, the number of drops, n, of a given size to kill a mosquito for any concentration of DDT in inert solvents is known. Now the problem is: What is the average number of drops, φt, that must strike an area so that there is a certain probability W_{n^+} that n or more drops strike any one area? This probability is the per cent mortality experienced by the mosquitoes when n drops contain the lethal amount M. Knowing the average φt necessary, the value of Ct_p is found immediately from equation 27.

Probability considerations: Assuming that the dispersion of the drops is random, a solution can be obtained by the method of Bateman (1). The probability that one drop strikes in time dt is φdt. Let $W_{n+1}(t + \mathrm{d}t)$ equal the probability that $n + 1$ drops strike in time $t + \mathrm{d}t$. This can occur in two ways: $n + 1$ drops in time t and none in dt, or n drops in time t and 1 in dt. This gives:

$$W_{n+1}(t + \mathrm{d}t) = (1 - \mathrm{d}t)W_{n+1}(t) + (\mathrm{d}t)W_n(t) \tag{28}$$

When dt approaches zero, we have:

$$\frac{\mathrm{d}W_{n+1}}{\mathrm{d}t} = (W_n - W_{n+1}) \tag{29}$$

The general solution of this set of equations is

$$W_n = e^{-\varphi t}\frac{(\varphi t)^n}{n!} \tag{30}$$

This is simply the Poisson distribution equation. W_n is the probability that exactly n drops will strike in time t if the average number per unit time is φ. However, we want to know the probability that n or more drops will strike. Calling this probability W_{n^+}, we have

$$W_{n^+} = e^{-\varphi t}\sum_{k=n}^{k=\infty}\frac{(\varphi t)^k}{k!}$$

or,

$$W_{n^+} = 1 - e^{-\varphi t}\sum_{k=0}^{k=n}\frac{(\varphi t)^k}{k!} \tag{31}$$

Equation 31 must be solved for a given value of n, which depends on the toxicity, the drop size, and the concentration of insecticide in the drops, to find the value of φt required to give each percentage kill. Three cases have been considered: $W_{n^+} = 0.50, 0.90$, and 0.99.

The calculations were carried out for small values of n by plotting W_{n^+} *vs.* φt for any given n and determining the value of φt necessary to give a chosen W_{n^+} graphically. For large values of n, i.e., for very small drops, the value of φt approaches the value of n, that is, the drops deposit in an almost uniform film.

Expressing d in microns and Ct_p in mg.min./cu.m., and taking $\sigma_h = 4.7 \times 10^{-2}$ sq. cm., the fundamental equation becomes

$$Ct_p = 6.16 \times 10^{-2}\, d\varphi t \text{ mg.min./cu.m.} \tag{32}$$

The results are shown in figure 5 for $M = 10^{-9}$ g. The curves exhibit a definite optimum drop diameter for each per cent mortality. The initial steep decrease is due to the rapid increase in settling velocity with increasing drop size, leading to a slope of -2 for 50 per cent mortality. For larger drops and larger percentage kills, the efficiency becomes smaller, resulting in a lower slope. Finally, at approximately the drop size where one drop contains the lethal dose, the minimum is reached. For larger drops the number of drops necessary to kill can no longer decrease, and the curve increases linearly with d.

Case II. Flying mosquito

For flying mosquitoes, in addition to the drops striking from above, there is a contribution to φ by impingement of the frontal area of the insect with the drops in the aerosol cloud. The fraction of drops which impact on the insect increases with drop size in accordance with Sell's theory. We then have

$$\varphi = \frac{gC\sigma_f}{3\pi\mu d} + \frac{6\epsilon C\sigma_v u}{\pi d^3\rho} \tag{32}$$

where u is now the speed of the insect through still air. The cloud concentration necessary to give this value of φ is

$$C = \frac{\varphi}{\frac{g\sigma_f}{3\pi\mu d} + \frac{6\epsilon\sigma_v u}{\pi d^3 \rho}} \tag{33}$$

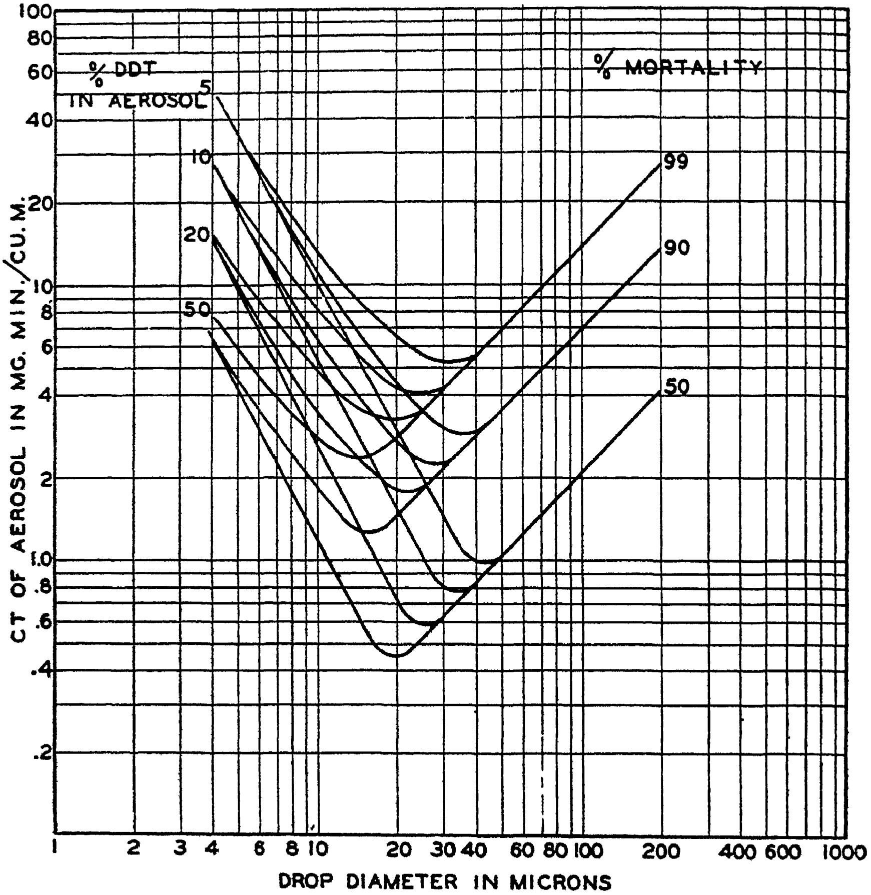


FIG. 5. Required *Ct* of aerosol for various mortality rates on flying mosquitoes. Lethal dose assumed to be 10^{-9} g.

ϵ is the fraction of the area dose deposited on the frontal area of the mosquito flying with a velocity u.

$$Ct_p = \frac{\varphi_t}{\frac{g\sigma_f}{3\pi\mu d} - \frac{6\epsilon\sigma_v u}{\pi d^3 \rho}} \tag{34}$$

ϵ is found from Sell's curves by assuming that we can represent the frontal area of the mosquito by a flat plate with an area of 3×10^{-2} sq. cm. σ_f is assumed to be 10^{-1} sq. cm.

If d is expressed in microns and Ct_p in mg. min./cu.m.,

$$Ct_v = \frac{1.67\varphi t d^3}{56.8d^2 + 2780\epsilon} \tag{35}$$

The value of φt was calculated as in case I. The results are shown in figure 6. The general shape of the curves is the same as in case I. There are two differences however. First, the Ct_p necessary to cause any percentage mortality in case II is approximately one-half that required in case I. This is mainly because the area σ_f is about twice the area σ_k. Secondly, the minima are shifted slightly to larger drop sizes in case II because of the fact that the ϵ increases with the square of the drop diameter.

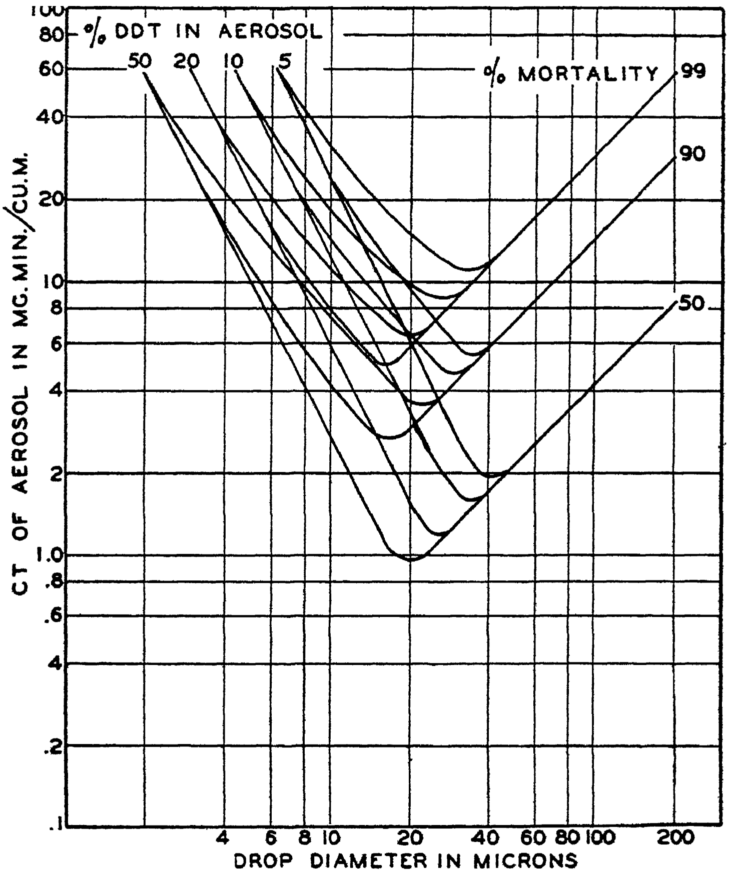


FIG. 6. Required Ct of aerosol for various mortality rates on resting mosquitoes. Lethal dose assumed to be 10^{-9} g.

Confirmation of the shape of the curves for 50 per cent mortality and the general position of the minimum may be found in the experimental data of La Mer and Hochberg, who studied the efficiency of DDT aerosols for killing mosquitoes in static atmospheres in a closed chamber (3). The insects were exposed to aerosols of uniform drop sizes containing 8 per cent insecticide in lubricating oil. The results, expressed as a median lethal dose for female *Aedes aegypti*, are shown in figure 7. The parallel lines represent the calculated

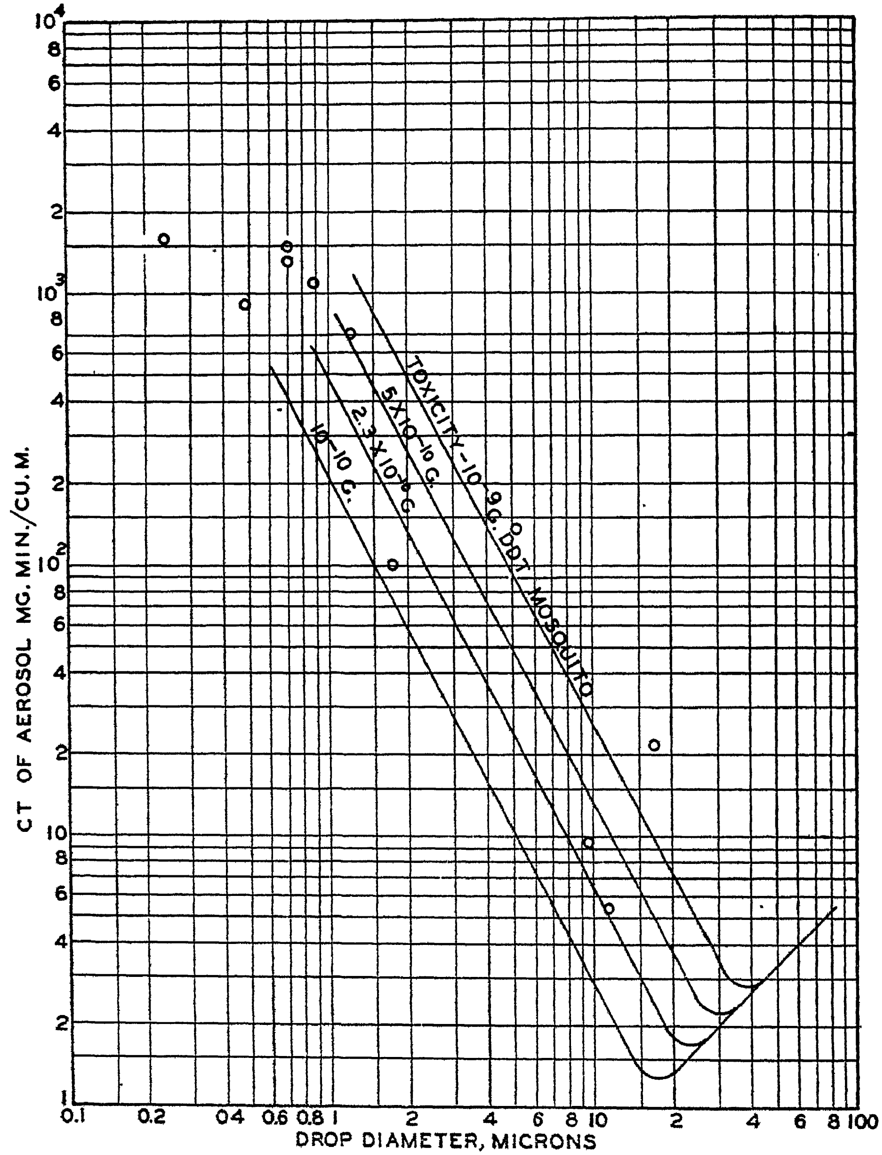


FIG. 7. Effect of particle size on dosage of DDT aerosol

effects for toxicities of 10^{-9}, 5×10^{-10}, 2.3×10^{-10}, and 10^{-10} g. of DDT per mosquito. The experimental data fall within this range. The calculated minimum dosage occurs at about 2 mg.min./cu.m. and at a drop diameter of about 30 microns for solutions of the given concentration. This is slightly beyond the range of drop size at which the experiments were made. In experiments of this nature it is difficult to prevent residue effects entirely when the insects rest on the screen.

Case III. Mosquitoes at rest on a screen in a moving air stream

Latta, La Mer, and coworkers have recently reported another series of experiments in which the effect of the drop size and wind velocity on the mortality rate was determined when mosquitoes were placed in a screen cage in a wind tunnel (4). It is interesting to compare the experimental results with those calculated from the probability equation, taking into consideration the fact that the efficiency of impaction of the droplets on the mosquito is also a function of the drop diameter and wind velocity. The authors noted that even at low wind velocities the insects did not fly but clung to the forward screen of the cage and remained stationary until the velocity reached about 16 miles per hour,

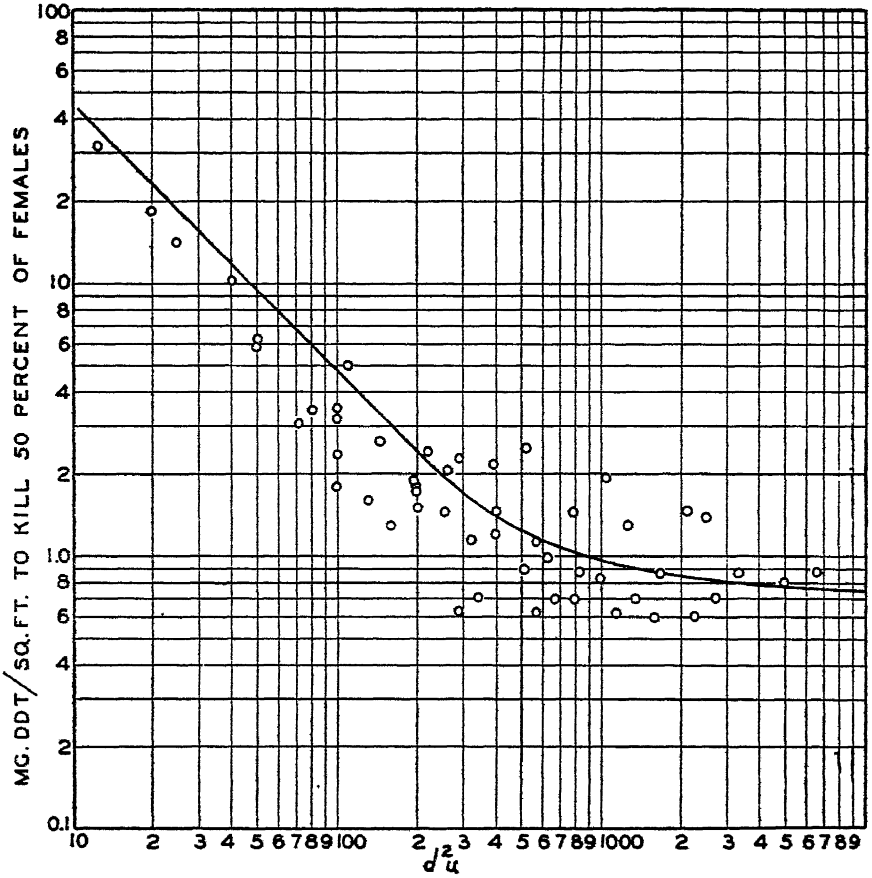


Fig. 8. Effect of particle size and velocity on amount of DDT aerosol required to produce 50 per cent mortality of mosquitoes. ———, calculated curve for toxic dose of 2.4×10^{-8} g. O, data of Latta *et al.* (J. Wash. Acad. Sci. 37, 397 (1947)).

when the force was sufficient to blow them to the opposite screen. Under these conditions the deposition takes place mainly by impaction, and the settling of the particles on the relatively small horizontal area may be neglected. The amount of aerosol deposited in unit time then becomes

$$\varphi = \frac{6\epsilon C u \sigma_h}{\pi d^3 \rho} \tag{36}$$

The concentration–time product is

$$Ct_p = \frac{\pi d^3 \rho \cdot \varphi t}{6 \epsilon u \sigma_h} \tag{37}$$

The impaction efficiency ϵ was calculated from the curve in figure 1 for a flat disk 3×10^{-2} sq. cm. in area.

From the values of φt calculated from equation 31 for $W_{n+} = 0.5$ the value of the aerosol dose in milligrams of DDT per square foot was calculated for various values of $d^2 u$. A value slightly lower than the median lethal value given by Latta and La Mer was used in these calculations: namely, $M = 2.4 \times 10^{-8}$ g. The density of the droplets was taken as 1 g. per cubic centimeter and the viscosity of air at 20°C. was used. The results of the calculation are shown by the curve in figure 8. The experimental data of Latta and La Mer are shown by the points. The agreement between the predicted and experimental data is excellent.

CONCLUSIONS

The agreement between Latta and La Mer's experimental results and those predicted from theoretical considerations lends support to the belief that the theory of aerosols may be applied to the dispersion of insecticides. Considerable discussion has taken place as to the relative merits of dispersing insecticide solutions as coarse drops and as aerosols of fine drops. There is theoretical and experimental evidence that small droplets are more effective than large drops when used under proper conditions. Practical difficulties of controlling the movement of aerosol clouds when dispersed from ground generators or from airplanes exist; yet, as the toxicity of insecticidal agents is increased, the advantages of using aerosols will also increase. The effectiveness of insecticidal aerosols dispersed from aircraft has been demonstrated in the field as a practical method of controlling natural insect populations. The problem remaining is that of working out the most effective technique of using the equipment on hand in order to secure the greatest economy of effort in use of the insecticide. Among the factors that are indicated in this study as important in aerial dispersion are such items as the type of aircraft, the position of the discharge on the plane, the speed of the plane, the height of the plane above the vegetation, the type of solvent used, the concentration of the solution, the quantity of solution dispersed, the drop spectrum and quantity of agent reaching the insect habitat, the type and height of vegetative cover, and the meteorological conditions.

It should be emphasized that neither theoretical studies nor laboratory experiments can be a complete substitute for field tests. Since the only object in the use of insecticides is to control insects, biological measurements must be the principal means of assessment of the results of any experiment. The importance of this has sometimes been overlooked in the eagerness of finding some physical measurement which could foretell the results. Good biological studies are admittedly difficult to make, especially when it is necessary to cover large areas and there is a natural fluctuation of the insect population due to weather conditions and breeding and flight habits. Careful correlation between the results on caged insects and natural population might eliminate some of this difficulty. The results on caged insects have often been misleading because of the deposition of the droplets on the screen and the slow rate at which equilibrium is obtained between the air in the cage and the conditions outside. Nevertheless, it should be possible to use caged insects to establish a definite correlation between biological results and the dosage and degree of contamination within the cage.

The helpful suggestions of C. W. Kearns, R. L. Le Tourneau, and Thomas Baron are acknowledged.

REFERENCES

(1) BATEMAN: Phil. Mag. **20**, 698 (1910).

(2) BOSANQUET, C. H., AND PEARSON, J. L.: Trans. Faraday Soc. **32**, 1249 (1936).

(3) LA MER, V. K., HOCHBERG, S., *et al.*: OSRD Report No. 4447 (December 11, 1945); J. Colloid Sci. **2**, 539 (1947); Chem. Revs. **44**, 341 (1949). Field tests are described by BRESCIA *et al.*: J. Econ. Entomol. **39**, 398 (1947).

(4) LATTA, R., ANDERSON, L. D., ROGERS, E. E., LA MER, V. K., HOCHBERG, S., LAUTERBACH, H., AND JOHNSON, I.: J. Wash. Acad. Sci. **37**, 397 (1947); OSRD Report No. 5566 (July 30, 1945).

(5) RODEBUSH, W. H.: OSRD Report No. 4757 (January 1, 1945).

(6) SELL, W.: Forsch. Gebiete Ingenieurw., Forschungsheft **No. 347** (1931).

(7) SUTTON, O. G.: Quart. J. Roy. Meteorol. Soc. **73**, 257 (1947).

(8) SUTTON, O. G.: Quart. J. Roy. Meteorol. Soc. **73**, 426 (1947).

ELECTRONIC METHODS OF COUNTING AEROSOL PARTICLES[1]

FRANK T. GUCKER, JR., AND CHESTER T. O'KONSKI[2]
Department of Chemistry, Indiana University, Bloomington, Indiana

Received October 18, 1948

Former methods of determining the particulate concentration in aerosols have been slow and laborious; but within the last few years two different rapid electronic methods of counting have been developed, the scope and range of which are discussed in this article. The first method depends upon light scattering, and the second upon electrostatic charging of the particles.

In the first method, developed by the authors and their collaborators, a fine stream of the aerosol, protected by a flowing sheath of pure air, passes through a spot under intense dark-field illumination and scatters flashes of light forward upon a photosensitive cell. Each particle about 0.6 micron or more in diameter causes an electrical pulse which is sufficiently large, after suitable amplification, to operate a mechanical counter. This apparatus will count particles weighing 5×10^{-13} g., at rates up to 1000 per minute.

A second type of electronic counter, developed by Arthur C. Guyton, depends upon the electrostatic charging of aerosol particles forced at high velocity through a fine jet to impinge upon a metallic collector. The electrical pulses imparted to the collector by particles of 2.5 microns or larger are amplified to operate a mechanical counter. Guyton found that the pulse amplitude is proportional to the square of the particulate diameter. This empirical fact we have explained in terms of the charging mechanism.

Both counters may be applied to determine the size of the aerosol particles by using suitable electrical discriminators. The electrostatic device may be very useful for larger particles, while improvements in the photoelectronic counter may make it applicable to particles even smaller than those counted at the present time.

INTRODUCTION

The determination of the particulate concentration of aerosols and other colloidal systems frequently is important for the colloidal chemist. Until recently, the only available techniques were slow and tedious, but two different electronic methods have been developed within the last few years to reduce the time and increase the convenience of these measurements.

---

[1] Presented at the Symposium on Aerosols which was held under the joint auspices of the Division of Physical and Inorganic Chemistry and the Division of Colloid Chemistry at the 113th National Meeting of the American Chemical Society, Chicago, Illinois, April 22, 1948.

The parts of this paper dealing with the researches of the authors and their collaborators are based upon work done for the Office of Scientific Research and Development under Contract OEMsr-282 with Northwestern University, and for Camp Detrick, Maryland, under Contracts WA-18-064-CWS-137 and -160 with Northwestern University and W-18-108-CM-31 with Indiana University.

For a fuller discussion of the design and operation of these counters than is presented here, see Chester T. O'Konski, "New Instrumental Methods in Aerosol Studies," Ph.D. Dissertation, Northwestern University, August, 1948.

[2] National Research Council Predoctoral Fellow in Chemistry at Northwestern University (1946–47) and Indiana University (1947–48). Present address: Department of Chemistry, University of California, Berkeley, California.

One of the earliest methods of counting aerosol particles involves the collection of a sample upon a suitable slide and a tedious visual count under the microscope. Several types of collector are available. One of these is an impactor, in which the aerosol passes through a fine jet at high speed and strikes a collecting plate to which it adheres. Efficient impactors have been described by May (19) and Sonkin (26). Another method of collecting aerosol particles depends upon their motion in the direction of a thermal gradient. This method was discovered by Aitken (1) and the theory of it has been discussed by Cawood (5). Recently the phenomenon has been studied by Rosenblatt and La Mer (24), who determined the velocity of thermal repulsion of individual small spherical particles of tricresyl phosphate. Comparison with the equations developed for radiometer theory indicated that thermal repulsion of aerosol particles is a radiometer phenomenon. This motion causes the dust-free space surrounding heated objects, investigated by Watson (29) and found to be clearly defined, extending to a distance approximately proportional to the square root of the thermal difference. Green and Watson (9) have described a dust sampler in which a wire is strung halfway between two closely spaced parallel cover glasses. The wire is heated so that the dust-free space is slightly greater than the distance between wire and cover glasses. When air is drawn across the wire, the dust precipitates just before it passes the wire. In another form of thermal precipitator, the sample of air is passed slowly between two closely spaced parallel plates differing in temperature by about 75°C., and the particles are deposited on the cooler one. The thermal precipitator is almost 100 per cent efficient.

A second counting method employs an ultramicroscope cell. The Zsigmondy slit ultramicroscope, in which the depth of the field is limited by the light beam, may give counts which are too high, because light scattered by the aerosol illuminates particles outside the desired volume. The effect increases with particle size. To eliminate this difficulty, Whytlaw-Gray and his collaborators (21, 22) developed a special cell of the cardiod type, in which the depth of field is limited by the walls of the cell. The apparatus is shown in figure 1. Light from an arc lamp G passes through a water-cooled slit E_1 and is focussed by means of a lens F upon a second slit E_3, where it passes through a microscope objective D. The light passing through the final slit E_2 provides uniform intense illumination of the space 0.1 mm. in depth between the parallel glass plates comprising the counting cell A. The viewing microscope C has a depth of focus greater than 0.1 mm., so that the depth of the cell limits that of the sample, while the area is limited by suitable stops in the eyepiece, chosen to give an average field of two or three particles. The inlet tube B is grounded and lined with moistened blotting paper to prevent loss of particles by electrification. The outlet tube is connected to suction through the stopcock H rotated by a wheel K connected to a motor. Thus successive samples of the aerosol are drawn in and stopped momentarily under the microscope. Only about a minute is required to count sixty fields, which contain enough particles to give an adequate statistical average. Since the field of view is 3×10^{-4} cc. or less, this counter is limited to aerosols containing about ten million particles per liter.

The walls of the cell are coated with a high-boiling liquid like *m*-xylene or paraffin oil, which wets glass and also wets any aerosol particles which strike the walls, rendering them invisible. Liquids as volatile as *m*-xylene also have another important effect. They condense upon the aerosol particles and brighten them appreciably. Whytlaw-Gray estimates that particles originally as small as 0.05 μ in radius may be loaded sufficiently to be counted, although otherwise they would be missed if they were below about 0.1 μ.

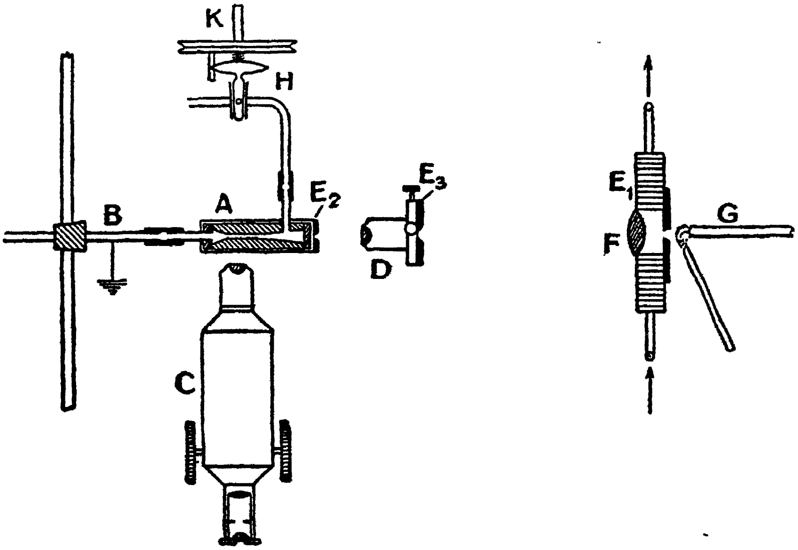


FIG. 1. Ultramicroscope counting apparatus of Whytlaw-Gray and his collaborators (from *Smoke* by R. Whytlaw-Gray and H. S. Patterson, Edward Arnold and Company, London 1932).

An extension of the condensation method of counting has been made by Green (8), who adapted the cloud chamber of Wilson (30) to this purpose. A small expansion chamber is connected to a standard Zsigmondy slit ultramicroscope cell and a cycle of operations is controlled by an electric motor. A sample of the aerosol is drawn into the cell, which is lined with damp filter paper. Then the valve is closed, the piston in the expansion cylinder suddenly drops, and an almost adiabatic expansion deposits moisture on the aerosol particles. The cell is photographed through the microscope, and then flushed out for another sample. A complete cycle takes only 2 sec., so that the counts are rapid. The usual difficulty with the Zsigmondy system is avoided in this case, because condensation builds up all the particles to nearly the same size; hence the volume under direct illumination is easily distinguished.

A third counting method employs bacteriological techniques. If an aerosol of single viable microörganisms is set up, they may be collected from a known volume of aerosol by means of a suitable cotton plug filter, or an impinger jet operating under water, as developed at Camp Detrick, and described by Rosebury (23). If the organisms are washed from the filter or collected in the impinger and deposited on agar, a colony count will give the number of particles. This process is limited in application and is very slow, ordinarily requiring about a day for incubation. It also involves the uncertainty of viability.

A PHOTOELECTRONIC PARTICLE COUNTER

During World War II gas-mask filters were improved to such an extent that they would remove all but a very small fraction of even the most penetrating aerosol particles (about 0.3 μ in diameter). The penetration of somewhat coarser test aerosols might be only a few particles per minute. In the summer of 1944 we undertook to develop an automatic instrument to count the individual particles in a rapidly moving aerosol stream and thus provide quantitative tests for the best available filters. By June, 1945, we had developed a photoelectronic apparatus which would count particles of dioctyl phthalate (DOP) down to about 0.6 μ diameter.

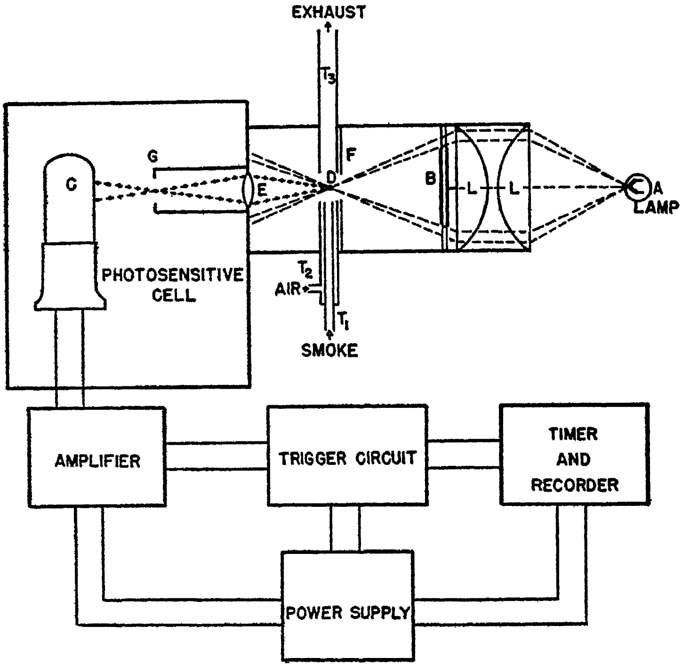


FIG. 2. Schematic diagram of photoelectronic particle counter

The general arrangement of the apparatus is shown in figure 2, while further details are given elsewhere (11). The lamp A is a 50-cp. automobile headlight, powered by a heavy-duty 6-v. storage battery. Inside the black-walled cell at D, the aspheric condenser lenses LL form a 4-mm. image of the filament of

A. Two circular black disks B cut out the center of the converging cone of light outlined by the dashed lines, thus providing dark-field illumination at D. The aerosol stream is introduced through the tube T_1, surrounded by a sheath of air flowing from T_2 at nearly the same linear rate. Thus each particle passes once through the intensely illuminated space at D, and is exhausted through T_3 without any chance of recirculating through the cell. Each particle scatters a flash of light, a part of which is collected by the lens E and focussed upon the photosensitive cell C, as indicated by the dotted lines. The electrical pulse from C is magnified up to 200,000 times in the amplifier, and then passed to a thyratron trigger circuit which actuates a mechanical recorder.

The optical system is like that developed in 1941–42 by La Mer, Sinclair, and Hochberg (16) for their study of aerosols carried out at Columbia University under contract with the National Defense Research Committee. It utilizes the light scattered in the near-forward direction, from about 5° to 30°. For particles with radii comparable to the wave length, a greater fraction of the light is scattered in this direction than through the same solid angle at 90°, as predicted by Mie (20) and verified by La Mer and Sinclair (16).

The blackened baffles F and G were designed to reduce background light in the system, and the black disk B inside the cell was found to prevent reflection from the glass plate of stray light which otherwise would find its way to the photosensitive cell C. The unit was built rigidly, and the lamp was held in a heavy clamp to prevent vibrations which vary the background light and may cause "optical microphonics" and spurious counts.

Two photosensitive cells were found useful in the counter. The first was the Thalofide cell, containing a special photoconductive surface of thallous sulfide, developed by Cashman (4) and his associates, working under a contract with the Office of Scientific Research and Development in the Physics Department of Northwestern University. High-resistance cells, with values of 5 megohms or more under background illumination, proved more sensitive than those of a lower resistance, which matched the 2.5-megohm input impedance of the amplifier more closely.

Later we found that a 931-A photomultiplier tube, operated at 30–50 v. per stage, could be used satisfactorily, after the background light had been reduced as much as possible. In this instrument, the stray light current in the 931-A tube was about ten times that due to an 0.8 μ DOP particle, while in the Thalofide cell the ratio was about 800 to 1, probably owing to long-wave stray light, to which this cell is more sensitive.

The electronic circuits of the counter required careful design and construction. The power supply was used to operate everything except the phototube, which was connected to a 22.5-v. dry battery, and the amplifier tube heaters, which were supplied by a 6-v. storage battery. The high-voltage supply was fed through a full-wave rectifier to two independent two-stage choke-input filters, one of which supplied the initial 1603 amplifier tube, chosen for its low microphonics, while the other supplied the two 6SJ7 amplifiers, trigger circuit, and recorder. This arrangement adequately reduced the rectifier ripple and prevented feedback from the second 6SJ7 to the 1603 tube.

The amplifier employed resistance–capacitance coupling between tubes, with wire-wound 5-megohm resistors in the input circuit from the phototube to reduce the noise, and the photocell and 1603-tube circuits enclosed in a desiccated brass box to decrease the possibility of excessive leakage currents and act as an electrostatic shield. The grid leads of the 6SJ7 tubes were made short, and shielded with a woven wire sleeve, soldered to the grounded chassis.

The biasing voltages in all three stages were furnished by cathode resistors, shunted by large condensers to extend the amplifier range down to about 10 cycles per second. This passed the low-frequency components of the non-sinusoidal photocell pulse and prevented pulse multiplication. The input resistors of the 6SJ7 tubes were shunted with suitable condensers to attenuate the high frequencies and improve the stability of the amplifier. This gives a response curve which is essentially flat from about 10 to 5000 cycles and falls to about one-half at 10,000 cycles. The gain of the amplifier was adjusted by means of potentiometers in the cathode circuits of the 6SJ7 tubes, which varied the negative feedback.

The trigger circuit employed a type-885 thyratron tube with a self-quenching choke-condenser plate circuit. The grid bias was controlled by means of a discriminator potentiometer, calibrated in volts with the zero reading at the firing point of the tube. The setting of the discriminator then determined the size of the minimum pulse which would fire the tube.

The time-recorder circuit contained two mechanical counters, which could be used in duplicate experiments, and a standard electric timing clock. Pressing a single button starts the timer and connects the trigger circuit to one of the counters. When the counter hand moves around to 100, it closes a contact in an automatic cut-off circuit which simultaneously stops both timer and counter. The operator then reads the time for 100 counts, which gives the same statistical uncertainty in each experiment. If a larger number of counts is desired, the automatic circuit is turned off, and then each revolution of the primary counter actuates an auxiliary *hundreds-counter*. At any time the automatic circuit can be turned on again to stop the instrument at the end of the next hundred counts.

In operation, the cell first is swept out with carefully filtered air, introduced through the outer tube at 3 l. per minute, until the dust count has fallen to one or less per minute. Two sheets of the best grade service-canister paper, or two complete pleated filters fastened into a canister, were found to give dust-free air *after they have been swept out for several hours or preferably overnight.*

In order to avoid using a flowmeter in the aerosol line, its flow is determined by difference between the effluent flow and the sheath-air flow, each read on calibrated flow meters. The standard rates are 4 and 3 l. per minute, making 1 l. per minute for the aerosol flow.

Careful tests of the counter were made to determine the size of particle and rate of counting for which it was applicable. Since the light scattering of an aerosol particle decreases rapidly with particle size, being proportional to the sixth power of the radius for very small particles, there will be a limit below which no particles will be counted. However, for dilute aerosols in which the particles are appreciably above this limit, each should register if the whole cross section of the stream at the focus is uniformly illuminated. As the concentration of the aerosol is increased some particles may be missed, owing to coincidences in the field of view and limitations of the photoelectronic circuits and mechanical counters.

Two methods were used to determine the upper limit of the original form of the counter. Both are described in the original publication (11), but only one will be treated here. The test aerosol was made up either of homogeneous spherical DOP particles of about 0.8 μ diameter, from a generator of the type developed by La Mer and Sinclair (15) in their OSRD work, or of spores of *Bacillus globigii* (BG), which are ellipsoidal, with a major axis averaging 1.2 μ and minor axes of about 0.8 μ, as shown by electron micrographs. The flow line using the BG aerosol is shown in figure 3. An aqueous suspension of 0.1 to 20 billion (10^9) BG spores per milliliter was sprayed through a nebulizer[3] into a chamber where it was mixed with enough dry air to form an aerosol of dry spores. This apparatus was kindly furnished by the group at Camp Detrick, Frederick, Maryland, with which we coöperated. They had found that this arrangement gave particles almost all of which were single spores. The number of viable spores in the suspension was determined at Camp Detrick. Despite the purification of the original material, there were always appreciable numbers of foreign particles from the nutrient medium.

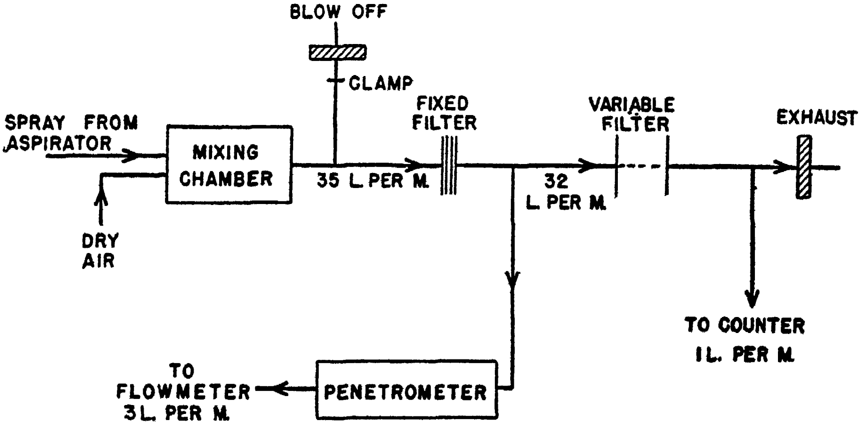


FIG. 3. Flow line for testing counter (from Journal of the American Chemical Society **69**, 2422 (1947)).

The aerosols were sent to the counter at a pressure somewhat above atmospheric, to avoid contamination with laboratory air, which often contains many thousands of particles per liter. The flow was adjusted by the variable blow-off to 35 l. per minute. The aerosol concentration was reduced by means of a fixed filter to a constant value, which was checked within 5 or 10 per cent by the reading of scattered light in the sensitive photoelectric photometer (penetrometer) developed by our group under an OSRD contract (10, 12). The dilute aerosol then was passed through a variable filter, made of sheets of paper, the transmission of each of which had been measured with the penetrometer, using a higher concentration of the same aerosol. A sample of the effluent aerosol was passed through the counter. From the observed counts per liter, c_2, and the known fractional transmission of the variable filter, P, the influent concentration was calculated as:

---

[3] Made by the Vapo-Nefrin Company, Upper Darby, Pennsylvania.

$$c_1 = c_2/P$$

The value of c_1 calculated thus was found to increase to a constant value c_0 at low counting rates. This value, divided by that at any higher rate, gives the corresponding count-rate factor, by which the observed count per minute should be multiplied to be consistent with the low-rate value. Figure 4 shows typical results with one of our early counters, obtained with a BG aerosol at Camp Detrick by Cohen and Voelker, on three successive days. Counts up to 1000 per minute were made with a count-rate factor of 1, but the correction increases greatly at higher rates. The practical range of the instrument is 3–1200 counts per minute.

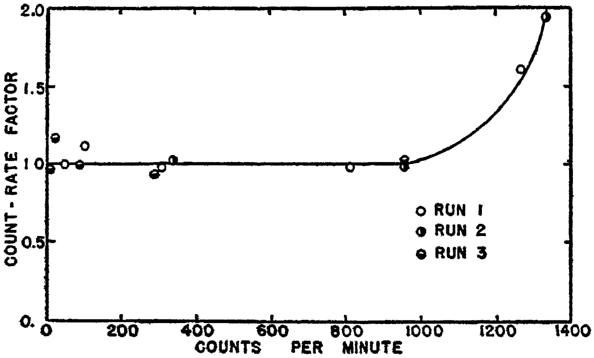


FIG. 4. Count-rate curve for BG aerosol, showing results obtained by Sgt. A. Cohen and Cpl. W. Voelker (from Journal of the American Chemical Society 69, 2422 (1947)).

The count-rate experiments did not determine whether all, or only a constant fraction, of the particles above a critical size are counted at low rates, and this is a difficult question to answer with certainty. Early experiments were carried out at Camp Detrick, in which the counts in our apparatus were compared with a colony count of viable spores collected on a cotton impinger, washed onto an agar plate, and incubated for 24 hr. These experiments showed more counts than viable spores, suggesting the presence of non-viable spores and extraneous matter from the cultural medium.

Later a series of experiments was carried out in which the concentration of viable spores reaching the counter was estimated from the concentration of the original aqueous suspension and quantitative dilution of the aerosol. The nebulizer was run with careful control of air pressure and flow rate. The BG suspension was diluted with water in five steps, to a ratio of 1/16, in order to vary the concentration of the aerosol considerably without changing the relative humidity, which might affect the scattering from the particles. The dilution steps were found to be self-consistent within ±8 per cent, as determined by the light scattered by the raw aerosols, measured in the penetrometer. The flow line was that shown in figure 3, except that the fixed filter was removed, and a three-stage diluting system was substituted for the variable filter, to avoid the selective removal of larger particles in a filter. The dilution factor of each stage was measured individually with the penetrometer, and the over-all transmission factor for the three stages used in these tests was calculated to be 0.030 per cent of the raw aerosol. Thus, the viable spores per liter in the diluted aerosol could be calculated, assuming no loss in the atomizer, and could be compared with the observed counts. Two counters were used. One, containing a Thalofide cell, was constructed in our Laboratory by Leo E. Farr for the use of Dr. Ronald M. Ferry of the Harvard Medical School. The other contained a 931-A photomultiplier. In these experiments, the count-rate factor increased from unity at about 500 counts per minute to 1.3 at 1000. At low rates, the average value of counts per viable spore in the atomized suspension was 0.3 in the Thalofide counter, and 0.5 in the photomultiplier counter. Later the lenses were adjusted to focus more red light on the aerosol stream, and values of 0.6 count per viable spore were obtained with the Thalofide cell. At first (11) we interpreted these results as due to lack of homogeneity in the 4-mm. image of the lamp filament, providing the dark-field illumination of the smoke stream, so that only a fraction of the particles were illuminated and counted. However, Dr. Ferry has pointed out to us recently that this difference may be due to the fact that a considerable portion of the original aqueous suspension fails to appear as aerosol particles in the output from the nebulizer.

Dr. Ferry has just completed a series of careful comparisons of the counts registered by our instrument and numbers calculated from the counts obtained in a two-stage impinger. He used both BG and *Serratia marcescens* (SM). The latter is slightly smaller, occurring in ellipsoidal or often nearly spherical form, 0.5 by 0.5 to 1.0 μ (3). In his counter, he found it necessary to use an 8-v. battery, giving 7 v. across the lamp bulb, in order to increase its intensity. He then found excellent agreement between the two methods, the ratio of counts to impinged specimens being 1.1 ± 0.1 for BG and slightly less for the smaller SM. These results have been given in more detail elsewhere in this Symposium (7).

SCOPE AND LIMITATIONS OF THE PHOTOELECTRONIC PARTICLE COUNTER

The photoelectronic counter was designed as a supersensitive penetration meter for testing gas-mask filters against aerosol particles about 1 micron in diameter. It is sensitive to single oil-smoke particles of 0.6 μ diameter, weighing less than 10^{-12} g. Spurious counts due to electrical noise in the system are kept below one per minute; hence when the counter is used at the present sampling rate of 1 l. per minute with a test aerosol containing 10^8 particles per liter, it will respond to penetrations of 1 micropercent, providing spurious counts due to dust are eliminated by blowing out the filter before the test. The upper range of the original counter is 1200 counts per minute; hence with the test aerosol of 10^8 particles per liter, the practical range of this instrument is about 3–1200 micropercent.

The lower limit of sensitivity is determined by the concentration of the test aerosol, the rate of sampling, i.e., the flow rate through the counter, the statistical errors which become significant when small numbers of particles are counted, and the dust particles blown from the filter under test. Since increasing the concentration of the test smoke increases clogging or aging effects in filters, this is not a satisfactory way to attain greater sensitivity for the very efficient filters. Increasing the flow rate through the counter also has its attendant problems. This cannot be done by increasing the cross section of the sampling tube greatly, unless a larger light source is used, since all particles must pass through the image of the light source in order to register. Another alternative, increasing the linear rate of flow, shortens the duration of the pulse of light upon the photosensitive cell. This is a disadvantage with both types of cell. Tests showed that the output signal voltage of the photoconductive Thalofide cell decreases approximately in proportion to the decreased length of the pulses of several milliseconds duration. Since the pulses at the present rate are only two to three times the largest fluctuations due to random electrical noise, the linear flow rate cannot be increased appreciably with this cell. The output signal voltage of the photoemissive electron-multiplier tube does not decrease with decreasing pulse length, but the amplifier soon becomes the limiting factor. For shorter pulses, the band width of the amplifier circuits must be increased in proportion to the increase in flow rate in order to maintain the same signal sensitivity. Since the noise increases as the square root of the band width, the signal-to-noise ratio will decrease as the square root of the flow rate. This again precludes use of much higher flow rates with the original counter.

Assuming that larger particles are being counted or that the signal-to-noise ratio can be increased in some other way, the above considerations show that a large increase in linear flow rate through the $\frac{1}{8}$ in. sampling tube of the counter might be tolerated in a counter using the electron-multiplier phototube. Another factor then enters,—namely, the laminar flow of the aerosol stream through the tubes in the smoke cell. At a certain limit turbulence might set in to make a further increase in flow rate impractical.

The upper limit on the counting rate of the present instrument is largely set by the resolving time of the thyratron counter circuit, which is about 10 millisec. This limitation can be overcome by using a scaling circuit with a shorter resolving time. However, after about a threefold increase in counting rate is achieved in this way, further increase is limited by the larger number of coincidences, caused by the presence of two or more particles in the field of view at the same time. According to statistical theory, the per cent of coincidences of randomly distributed aerosol particles in a given illuminated volume of sample goes up linearly with the concentration. This volume coincidence effect is independent of the flow rate through the cell when considered as an upper limit on the *concentration* of smokes which can be counted. Decreasing the cross-sectional area of the sampling tube will raise this limit proportionally without introducing further difficulties. However, decreasing the length of the flow path over which the particle is illuminated and viewed will have the same effect as increasing the linear flow rate, in causing a decrease in signal-to-noise ratio, as explained in the preceding paragraph.

The lower limit of particle size is likewise determined by the signal-to-noise ratio in the instrument. If this can be improved, smaller particles can be detected. The signal amplitude can be increased by using a more intense light source or by improving the optical system to collect and focus sharply a greater amount of light on the particle and to collect more of the light scattered by the particle. The amount of stray light in the optical system is important for two reasons: (*1*) it determines the output current and noise level in the electron-multiplier phototube, and (*2*) if it is as large as in the original instrument, relatively small variations in the lamp and vibrations of the system may produce spurious pulses. For both reasons, the stray light in the optical system should be decreased as much as possible, preferably without decreasing the signal amplitude for smaller particles.

A new optical system, under consideration for some time, has been constructed and is now being tested. Preliminary results show that the stray light in the system has been reduced so much that the background-light current in the photomultiplier tube is about the value of the dark current. In addition, a differential pulse amplitude selector circuit has been designed and constructed. This will pass only the pulses in a given range of amplitude, and will allow automatic discrimination against the larger pulses from dust particles. Tests are being continued and the improvements and extended applicability of the apparatus are being evaluated according to the considerations discussed above.

APPLICATIONS OF PHOTOELECTRONIC PARTICLE COUNTING

In addition to its original use as a penetration meter to test the best gas-mask filters which can be produced, the photoelectronic counter has many possible applications in other fields.

The counter built for Dr. Ferry is now being studied by his group at Harvard University to evaluate its usefulness in counting bacterial aerosols (7). Other bacteriological applications may be made. It is very convenient in testing for leaks which might bring contamination to a system where sterile conditions are desired. A trace of unfiltered room air will instantly register on the counter, while the accompanying bacterial contamination would require many hours to show up in the usual plate counts. The instrument can be used to count various particles suspended in liquids either by adapting the optical system to liquids, or by atomizing the liquid suspension and counting the particles as aerosols. Particles which differ appreciably in size could be distinguished and counted separately by several settings of the *discriminator*.

The improved counter now under study contains a pulse-amplitude selector circuit which will permit the counting of pulses within a chosen range of voltage, and may be applied to a rapid determination of particle-size distribution, at least in spherical liquid aerosol particles, where the light scattering can be calculated by means of the Mie theory (16, 20).

AN ELECTROSTATIC PARTICLE COUNTER

Shortly after the development of our photoelectronic particle counter, Guyton (13) developed at Camp Detrick an electrostatic apparatus which can count aerosol particles with a diameter of 3 μ or more. His detector unit consisted of a tube, presumably of glass, tapering at an angle of 45° to a hole 0.8 mm. in diameter, with a collector directly in front of the orifice. This collector was either a metal plate or a copper wire, 0.4 mm. in diameter. The latter was found to be more sensitive and was used in most of Guyton's work. By applying a vacuum beyond the tapered orifice, aerosol particles are drawn through the tube to impinge upon the copper wire, which is connected to a high-frequency amplifier with resistance and choke coupling to give a response above 10,000 cycles per second and an amplification of about 100,000. Individual particles striking the wire produce impulses which are amplified in the first three stages and rectified by the grid of the fourth stage of amplification. This serves to lengthen the pulses from about 50 to 500 microsec., determined by the *RC* value of the resistance-capacitance coupled grid circuit. The pulse then fires a thyratron tube, whose discharge in turn is amplified through a class C triode which energizes a mechanical counter. This instrument was used at counting rates up to about 1800 per minute, or 300 per liter, since the flow rate through the orifice was about 6 l. per minute.

Guyton stated that the various solid aerosol particles he tested all produced on the wire positive pulses which increased in voltage with the square of the particle diameter, or the surface area of the particle. The polarity and size of the pulses he found to be independent of the potential at which the detector wire was maintained. Water droplets behaved very differently. They gave positive pulses with the wire at ground potential, no pulses with the wire at plus 22 v., and negative pulses increasing linearly with the wire potential beyond this value. He concluded that the "surface voltage on water particles is plus 22 volts" and that the "natural electrostatic charge" of water particles can be reversed by an applied voltage.

Guyton was doubtful about the mechanism by which the electrostatic charge is produced, but concluded from the linear increase of the pulse size with the surface area of the particle that the charge may be produced by friction between the particle and turbulent air in the nozzle or between the particle and the wire. He indicated that the pulse duration of 50 microsec. could be decreased to 10 microsec. or less by decreasing the *RC* value in the electrical connections to the hookup wire, and that the theoretical upper limit of the counting rate would therefore be about 10^5 per second. He also concluded that the lower limit of particle size depended upon the particular amplifier he used, which could record only pulses of 60 microvolts or more, and that the design of the amplifier could be improved "to detect impulses of a few microvolts which correspond to particles of 0.3 to 1 micron in size."

The electrostatic charging of dielectrics, or so-called frictional electricity, has been observed and studied for a long time. Most of the studies have been empirical. Perhaps the first utilization of the electrostatic charging of small particles was that by Armstrong (2), who constructed a steam generator and jet which he used as an electrostatic generator as early as 1843. Subsequently Michael Faraday (6) undertook a study of the mechanism by which the charge was produced. His ingenious experimentation with limited equipment led to many interesting observations and speculations, including the notion that the charging was due to a rubbing or frictional process. Since his time, a great deal of work has been done on electrostatic properties and a review of the phenomena in general by Loeb (17) shows "a most unsatisfactory array of discordant, often non-reproducible and confusing results, whose theoretical interpretation in some cases is quite hopeless."

Whatever the charging process, there is evidently a theoretical upper limit to the amount of charge which can exist on a small particle. If the charge on an aerosol particle is increased, a point is reached where the potential gradient at the surface becomes sufficient to produce breakdown in the surrounding air, which would tend to discharge the particle and prevent further accumulation of electricity. Hence for a spherical particle of radius r the maximum possible charge q can be calculated by making use of the relation $q = CV$, where C is the capacity of the particle, which in electrostatic units equals the radius in centimeters, and V is the potential in volts. Taking the derivative we get for the change of potential with drop size the expression: $\mathrm{d}V/\mathrm{d}r = -q/r^2$. This may be shown to be identical with the expression for the potential gradient at the surface, which can be set equal to the breakdown potential gradient in air, about 30,000 v. per centimeter or 100 E.S.U. per centimeter. This gives, for the maximum charge, a value of $100\ r^2$ E.S.U., where r is in centimeters, or $3.3 \times 10^{-8} r^2$ coulomb. Hence we observe that the maximum permissible charge on the particle is proportional to the square of the radius or the surface area of the particle. This relation has been used by Vollrath (28) in a description of his high-voltage electrostatic generator, employing finely divided diatomaceous earth blown through copper tubes which become highly charged in the process. Airplanes may build up a high charge as they fly through dust, clouds, rain, or snow, and this causes corona discharges—St. Elmo's fire—which interfere with radio reception. Hall (14) studied the possibility of discharging airplanes by blowing diatomaceous earth through brass tubes filled with metal turnings. He evidently used the same equation for the maximum charge per particle, to show that the maximum amount of charge which could be transferred by 1 g. of finely divided material varies inversely with the radius of the particles. For nonspherical surfaces the same considerations apply and Silsbee (25) computes a limiting charge density of 2.65×10^{-9} coulomb per square centimeter, which is in agreement with the result of our calculations.

SUGGESTED MECHANISM OF THE ELECTROSTATIC COUNTER

Although several mechanisms can produce electrostatic charging effects, as has been indicated by Loeb (17), the authors prefer to explain the action of Guyton's electrostatic counter by means of the mechanism suggested by Vollrath (28) in accord with the views of Helmholtz on so-called frictional electricity. This postulates that a double layer of charge is set up at the surface of contact between the particle and the wire. Because of the small distance between the surfaces (10^{-8} to 10^{-7} cm.) at contact, the effective capacitance is high and a low contact potential can result in the transfer of a substantial quantity of charge according to the relation $q = CV$. After a charge transfer has occurred, the surfaces become separated. As the particle moves away from the wire, its capacity with respect to its surroundings rapidly decreases with distance from the wire. The charge q, trapped on the particle, will cause the potential V to rise accordingly. If the potential gradient at the surface exceeds the breakdown potential of air, ionization of the surrounding air will occur and the excess charge will be lost. This most probably would occur when the particle is still very close to the wire, so it appears reasonable that the excess might be largely returned to the wire, resulting in a net charge transfer equal to the maximum permissible charge on the particle. In this event, the voltage of the pulses on the wire can be calculated directly from the size of the particle and the capacity of the input circuit to ground:

$$V_p = q_{\max}/C_i = 3.3 \times 10^{-8}r^2/C_i$$

where r is the radius of the particle in centimeters, C_i is the input capacity in farads, and V_p is the pulse amplitude in volts. Converting to microns, micromicrofarads, and microvolts, we obtain the relation $v_p = 330r^2/c_i$.

INTERPRETATION OF GUYTON'S RESULTS

The preceding equation may explain why Guyton found the wire collector more sensitive than the metal plate, which would be expected to have a higher capacity to ground, increasing the value of c_i.

The empirical relation which Guyton found for particle diameter d in microns as a function of pulse size v_p in microvolts was:

$$d = v_p^{1/2}/3.04$$

Comparing this with our equation we obtain a value of 9 micromicrofarads for the capacity of the input circuit. The rated input capacity of the first tube in Guyton's amplifier, a 1603, is 4.6 micromicrofarads. The difference between these two values could readily be due to the distributed capacity of the detector wire, the blocking condenser, and the grid leads.

The capacity of the input circuit can also be estimated from the duration of the pulses, which Guyton placed at about 50 microsec., and the resistance to ground of the equivalent input circuit, 1.5 megohms. Setting RC equal to the pulse duration, this yields the approximate result of 33 micromicrofarads which also might be possible, depending on the physical arrangement of the grid blocking condenser and the leads in the amplifier. However, this gives in our equation

$v_p = 10r^2$. This is not consistent with Guyton's empirical relation, which can be written $v_p = 37r^2$. The visual estimation of a short pulse on an oscilloscope could be considerably in error, and could yield a result far from the true RC value of the input circuit. Hence further investigations are required to establish whether or not the assumptions introduced above are valid.

To account for a greater charge transfer than expected on the basis of the above postulates, two points merit further investigation: (*1*) the chance that a particle may be charged before striking the wire and (*2*) the possibility that small particles may carry away a greater charge than they could be expected to retain under normal conditions. The first effect might double the charge transfer. If, for example, before a particle came in contact with the wire, it were charged positively to the maximum value given by $q = 3.3 \times 10^{-8}r^2$, perhaps by contact with the orifice wall, it could effect a charge transfer of $6.6 \times 10^{-8}r^2$, providing it left the wire with the maximum negative charge. In connection with the second effect, the processes involving ionization in air, corona discharges, etc., are very complex and subject to statistical variations which are often difficult to interpret, as pointed out by Loeb and Kip (18). The electrostatic detector may prove a useful tool for investigations of these phenomena.

SCOPE AND APPLICABILITY OF THE ELECTROSTATIC COUNTER

In its original form the electrostatic detector can count only particles with a diameter of about 3 microns or more. The above discussion clearly indicates that the lower limit of sensitivity might be extended to somewhat smaller particles, perhaps 1 micron or more in diameter, by improved design of the amplifier. A value of about 5 micromicrofarads for the input capacity is feasible with existing electronic equipment.

Another important consideration must be thoroughly investigated, namely, the input noise of the amplifier circuit. The root-mean-square value of the thermal noise in an amplifier is given (27) by the equation:

$$= \sqrt{1.64 \times 10^{-20} R\Delta f}$$

where R is the input resistance in ohms, and Δf is the band width in cycles per second. Thus the RMS thermal noise alone in an amplifier with a band width of 10,000 cycles connected across a 1.5-megohm load is about 16 microvolts. Random peak fluctuations can easily exceed two or three times this value. It therefore appears that little can be gained by increasing the amplification of the amplifier used in the electrostatic counter because, above a certain level, random noise pulses will give spurious counts. Any attempts to increase the counting rate to the proposed value of 10^5 per second will require an amplifier of greater band width and will therefore decrease the sensitivity of the unit.

To investigate the mechanism and the possible improvements in the electrostatic counter, and to compare it with the photoelectronic counter, several detector units and the necessary amplifier and counter circuits have been set up in our laboratory. For larger particles this type of counter might prove useful in discriminating between particles of different chemical compositions, and a correlation between results obtained with it and our photoelectronic counter might yield additional information on the aerosol system under test.

REFERENCES

(1) Aitken, J.: *Collected Scientific Papers*, p. 84. Cambridge University Press, London (1923).
(2) Armstrong, W. G.: Phil. Mag. **23**, 194 (1843).
(3) Breed, R. S., and others: Bergey's *Manual of Determinative Bacteriology*, 6th edition, p. 480. The Williams & Wilkins Company, Baltimore, Maryland (1948).
(4) Cashman, R. J.: "Development of Stable Thallous Sulfide Photoconductive Cells for Detection of Infra-red Radiation," OSRD Report No. 5997; J. Optical Soc. Am. **36**, 356A (1946); Proceedings of the National Electronics Conference **2**, 171 (1946).
(5) Cawood, W.: Trans. Faraday Soc. **32**, 1068 (1936).
(6) Faraday, Michael: *Experimental Researches in Electricity*, Vol. III, pp. 106–26. Bernard Quaritch, London (1844).
(7) Ferry, R. N., Farr, L. E., Jr., and Hartman, Mary G.: Chem. Revs. **44**, 389 (1949).
(8) Green, H. L.: Phil. Mag. [7] **4**, 1046 (1927).
(9) Green, H. L., and Watson, H. H.: Medical Research Council, Special Report Series, No. 199, His Majesty's Stationery Office (1935).
(10) Gucker, F. T., Jr.: Electronics **20**, 107 (1947).
(11) Gucker, F. T., Jr., O'Konski, C. T., Pickard, H. B., and Pitts, J. N., Jr.: J. Am. Chem. Soc. **69**, 2422 (1947).
(12) Gucker, F. T., Jr., Pickard, F. B., and O'Konski, C. T.: J. Am. Chem. Soc. **69**, 429 (1947).
(13) Guyton, A. C.: J. Ind. Hyg. Toxicol. **28**, 133 (1946).
(14) Hall, W. C.: J. Applied Phys. **18**, 759 (1947).
(15) La Mer, V. K., and Sinclair, D.: NDRC, Division B, Report No. 57 (OSRD Report No. 119 (July 15, 1941); OSRD Report No. 1668 (August 20, 1943); Chem. Revs. **44**, 245 (1949).
(16) La Mer, V. K., and Sinclair, D.: OSRD Report No. 1857 (1943), available through Office of Publications Board, U. S. Department of Commerce, as Report No. 944; also Part V of OSRD Report No. 865 (September 4, 1942), "A High Sensitivity Smoke Meter" (D. Sinclair), and "A Balanced Photoelectric Smoke Filter Penetration Photometer" (S. Hochberg).
(17) Loeb, L. B.: Science **102**, 573 (1945).
(18) Loeb, L. B., and Kip, A. F.: J. Applied Phys. **10**, 142 (1939).
(19) May, R. K.: J. Sci. Instruments **22**, 187 (1945).
(20) Mie, G.: Ann. Physik **25**, 377 (1908).
(21) Nonhebel, G., Colvin, J., Patterson, H. S., and Whytlaw-Gray, R.: Proc. Roy. Soc. (London) **A116**, 540 (1927).
(22) Patterson, H. S., Whytlaw-Gray, R., and Cawood, W.: Proc. Roy. Soc. (London) **A124**, 502 (1929).
(23) Rosebury, T.: *Experimental Air-Borne Infection*, p. 108 *et seq.* The Williams & Wilkins Company, Baltimore, Maryland (1927).
(24) Rosenblatt, P., and La Mer, V. K.: Phys. Rev. **70**, 385 (1946).
(25) Silsbee, F. B.: *Static Electricity*, Bureau of Standards Circular C438, U. S. Department of Commerce (1942).
(26) Sonkin, L. S.: J. Ind. Hyg. Toxicol. **28**, 269 (1946).
(27) Strong, J.: *Procedures in Experimental Physics*, p. 436. Prentice-Hall, Inc., New York (1941).
(28) Vollrath, R. E.: Phys. Rev. **42**, 298 (1932).
(29) Watson, H. H.: Trans. Faraday Soc. **32**, 1073 (1936).
(30) Wilson, C. T. R.: Proc. Roy. Soc. (London) **A87**, 277 (1912).

THE PREPARATION AND MEASUREMENT OF THE CONCENTRATION OF DILUTE BACTERIAL AEROSOLS[1, 2]

RONALD M. FERRY, LEO E. FARR, JR., AND MARY G. HARTMAN

Department of Bacteriology and Immunology, Harvard Medical School, Boston, Massachusetts

Received October 18, 1948

Since the suspicion that disease may be airborne is centuries old, it is not surprising that medical scientists have long been concerned with the properties of airborne particulate suspensions. Following the rise of modern bacteriology, the medical significance of dilute aerosols has become clear. And, since many bacteria may be cultured to form particles with nearly uniform size and shape, characteristic of a given species, aerosols derived from bacterial suspensions may be profitably studied. Knowledge so gained should extend the understanding of colloidal phenomena and also clarify the mechanism of airborne disease.

Quantitative studies surely depend on, first, methods for generating uniform dilute aerosols, second, accurate measurement of their concentration and, third, with living organisms, on their collection and enumeration by biological methods, and further understanding of the physical properties of particulates. Although methods of generation are not completely satisfactory, concentration has been determined by the method of impingement and direct photoelectronic count. The latter method is convenient, rapid, and accurate provided particle size somewhat exceeds 0.7μ. Although impingement yields results consistent with the dimensionless group,

$$\frac{\rho V D^2}{\eta l}$$

derived by May and Sell, a comprehensive theory must include such factors as particulate asymmetry, the physical characteristics of the particulate and collecting surfaces, and the microaerodynamics of the impinger used. Biological enumeration is not only slow, but apparently limited by the efficiency of impingement methods. Simultaneous application of photoelectronic and biological methods should help increase understanding of aerosols composed of solid particles.

INTRODUCTORY REVIEW

Since disease has long been associated with airborne agents, it is not surprising that medical scientists have long concerned themselves, implicitly or explicitly, with the study of aerosols. Indeed, the suspicion that infection may be airborne is probably as old as medicine itself. In *The Evolution of Modern Medicine*, Osler (66) cites Maspero (59) as his authority for stating that "the Egyptians believed that disease and death were not natural and inevitable, but were caused by some malign influence which could use any agency, natural or invisible, and

---

[1] Presented at the Symposium on Aerosols which was held under the joint auspices of the Division of Physical and Inorganic Chemistry and the Division of Colloid Chemistry at the 113th National Meeting of the American Chemical Society, Chicago, Illinois, April 22, 1948.

[2] This paper is based upon work sponsored in part by the Biological Department, Chemical Corps, Camp Detrick, Frederick, Maryland, under Contract No. W18-064-CWS-151 with Harvard University.

very often belonged to the invisible world." Fifty or sixty centuries were to pass before the nature of some of the invisible agencies began to be apprehended. Our understanding is, of course, to be attributed to the rise of bacteriology and the concomitant development of appropriate investigative tools late in the nineteenth and early in the twentieth century. More recently virology has further clarified our understanding of many other diseases. And it is now proven that the air we breathe is a disease vector.

None the less, the importance of airborne factors has been obscured until recently by at least two factors. In the first place, Manson (58) in 1879 demonstrated that mosquitoes were responsible for the spread of filariasis. Subsequently, insect vectors have been associated with such diseases as malaria, plague, and typhus. Second, it has been technically more difficult to obtain evidence directly connecting airborne particles with disease. For although airborne dust particles are ubiquitous and generally numerous, their sizes range from a small fraction of a micron upwards and completely efficient collection is difficult. Moreover, only a small proportion of those particles may be expected to contain an infectious agent.

None the less, epidemiological investigations of the more common respiratory diseases such as pulmonary tuberculosis (1), influenza, and the common cold demonstrate the importance of airborne agents (16). On this basis, air sterilization by ultraviolet light (27, 57) or by chemical agents (4, 5, 15, 26, 28, 29, 31, 42, 54, 60, 70, 71, 72, 74, 75, 87, 88) has been applied to their control. These methods will not be reviewed here. Nor have laboratory studies of airborne infection been neglected (1, 3, 23, 24, 26, 32, 55, 68, 75, 77, 80, 86, 89, 93, 94). These studies have generally been based on exposure of animals to aerosols generated by the atomization of living suspensions of the infectious agent. The concentration of the bacterial aerosol was usually estimated from samples collected either in liquids or semisolids by means of a modified Greenburg-Smith (77) impinger, a cotton filter, or by means of the Wells air centrifuge (92). The estimate of the number of organisms contained in the sample rested primarily on counts of the number of bacterial colonies appearing in plate cultures of an aliquot.

Although the results obtained by these methods are extremely valuable, the methods are time consuming, often inconvenient, and affected by the many variables encountered in biological systems. Moreover, they provide no possibility for direct estimation of the ratio of living to dead organisms in the aerosols studied. And this information appears necessary for an understanding of the factors affecting viability and consequent infectivity. Despite the recent contributions of Rosebury (77) and Robertson (74) and his coworkers, further information seems desirable.

Injury to the respiratory tract and respiratory infection have been related to invasion of the respiratory tract (11, 12, 17, 18, 20, 30, 43, 50, 52, 85, 95). Rooks (76) has reported a study of bacterial penetration. Further studies relating penetration not only to the size and dispersion but to other physical properties of the particles appear desirable. Recently Wilson and La Mer (97) have reported a study with radioactive tracers.

Moreover, information concerning the number of bacteria necessary to produce infection in man and animals would appear to depend on rapid and accurate methods of determining the particulate count in rather dilute aerosols.

Finally, certain bacterial species may be grown under conditions such that their size and shape are nearly uniform, although characteristic of each species. Since size varies about twentyfold from 0.4 to 8.0 μ, and the shape from the almost spherical cocci to the thin rodlike bacilli, it is possible to secure valuable physical information concerning aerosols containing particles differing not only in size but also shape, from a study of airborne bacteria. Unlike most solid particles, their density is low.

As a result of these considerations, it seemed to us that quantitative laboratory studies depended on: (*1*) a method for the preparation of pure bacterial aerosols at low concentrations, from one hundred to a few thousand per liter, and preferably containing only single cells of nearly uniform size; (*2*) a rapid and accurate method for determining the number of particles in such suspensions; and (*3*) a satisfactory method for collecting and enumerating the number of living bacteria in aerosols. We believed that the first requirement could readily be met by slight modification of existing methods for atomizing liquid suspensions of particles. We hoped that the photoelectronic counter, which utilizes a lens system developed by La Mer and Sinclair (47) and a photoelectronic, amplifying, and counting system developed by Gucker (38), would satisfy the second. And we expected that these methods combined with the use of one of the devices frequently used for collecting airborne organisms would satisfy the third requirement. This paper describes the first and appraises the second method, and in partly appraising the third, suggests that further scrutiny of the methods may be necessary. For convenience, the methods are treated in separate sections. In their present form, they are applicable only to laboratory studies and are not yet sufficiently refined to permit the study of individual particles much smaller than 1.0 μ.

Since Gucker's photoelectronic counter has been extensively used by us, and is often referred to, it is briefly described here. It is more fully described in a preceding paper (39). This instrument depends primarily on light scattering by small particles suspended in air (83, 90). The apparatus is schematically shown in figure 1, to parts of which reference is made by letter.

By means of appropriate lenses (A, B) and stop system (C), a hollow cone of light is brought to a rather deep focus, through which an air stream may be passed. This part of the apparatus we shall subsequently call the smoke cell. Unless the air contains particles, the rays pass outside a small collecting lens (E). If suspended particles are present, light is scattered with maximal intensity in the forward direction. Some of the scattered light enters lens D and is brought to a focus at diaphragm E. It then excites a photocell (F); the resulting impulses are amplified by an appropriate electronic system and may be electronically or mechanically recorded. For further details, the reader is referred to Gucker's papers (38, 39).

I. GENERATION OF DILUTE BACTERIAL AEROSOLS

A. Introduction

Accurate studies of aerosols appear to depend on the production of homogeneous suspensions of uniformly sized single particles in air. Although the treatment of spherical particles is simpler, the theory of light scattering is sufficiently developed for the consideration of less symmetrical particles (7, 19, 45, 47, 62, 65, 69, 81, 98). Since the Gucker instrument with a Thalofide photocell responds accurately only if the counting rate is less than 1000 per minute (38), dilute aerosols are essential for study.

Bacterial aerosols can be readily generated (13, 74, 77, 88), and, provided the test organisms are suitably selected, reasonable uniformity of particle size is attainable. Perfect spherical symmetry is not readily realized, for although cocci are nearly spherical, they grow in clumps which are not readily broken. Consequently, the resistant spores of *B. globigii* and the sensitive vegetative form *S. marcescens* were chosen. These organisms have similar symmetry, and appear as ellipsoids of revolution with axes in the former case about 0.8 × 1.2 μ (8), in the latter approximately 0.6 × 0.8 μ (8). Microscopic examination showed little variation in the individual cells of each species. The difference in size seemed suitable for fixing the limit below which the photoelectronic instrument fails accurately to estimate concentration. Moreover, penetration of the respiratory tract, which depends on particle size and dispersion (11, 12, 43, 50, 95), is near a maximum in this size range, and our studies have a biological as well as physical orientation. Finally, the difference in size appeared to permit theoretical analysis.

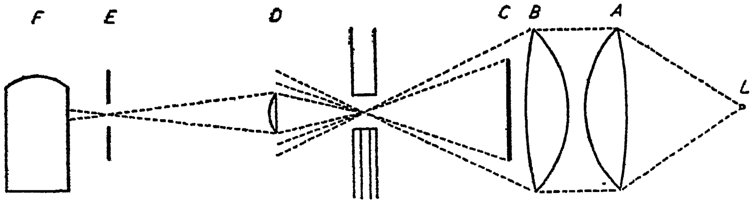


FIG. 1. Schematic representation of optical system for photoelectronic counter

In generating bacterial aerosols, the following precautions should be observed: (*1*) Leakage inwards or outwards is not permissible; the former will contaminate the aerosols, the latter will pollute room air. (*2*) Humidity must be controlled and readily measurable, since the death rate of bacteria apparently has a critical maximum at about 50 per cent relative humidity (25). (*3*) Diluting air must be carefully filtered, not only to avoid bacterial contamination, but also to eliminate adventitious photoelectronic counts. Future quantitative studies, especially with live infectious agents, depend on observing all of these precautions. For many purposes, this has not been necessary.

Simple atomization of bacterial suspensions either in water, broth, or other media, followed by simple dilution in a chamber (74, 77, 87, 93, 94), has generally sufficed. The atomizers used may be classified as refluxing or direct. Both depend on the shearing effect of a high-velocity air stream which is directed across an orifice through which the liquid passes and is broken into droplets. In the former, only smaller droplets and particles escape (34, 77, 87), while larger droplets return to the reservoir. Consequently some of the fluid is repeatedly atomized so that partial evaporation of the larger droplets within the atomizer causes concentration of the suspension atomized and consequent concentration of the aerosol. In the direct type concentration is avoided, but droplets are larger (87). With either type the observed aerosol concentration is less than that calculated from air flow and the volume and concentration of the suspension atomized (77, 87). This effect is much greater with refluxing than with direct atomizers and may be attributed to impingement of incompletely dried droplets on nearby surfaces (13). Further attenuation is readily observed in regions of turbulence, as in mixing chambers.

The recovery of particles in an aerosol prepared with a given atomizing and diluting system is, under controlled conditions, almost constant. For example, in our experience with atomizer 8 and a *B. globigii* suspension, recovery was about 35 per cent of that expected. The liquid orifice may, however, unexpectedly become occluded so that constant control of particulate concentration at the point of use is essential.

After careful consideration, the refluxing type was deliberately chosen for use in our studies, since it had been reported to produce aerosols composed chiefly of single bacterial cells, presumably by breaking up bacterial clumps by repeated reatomization. For our purposes, uniform particulate size was more important than constant concentration. Further improvements in atomizer design should make it possible to maintain more nearly constant concentration of aerosols.

B. Methods

A general description of the apparatus used is followed by a brief description of certain important procedures.

Atomizer: Readily available Vaponefrin atomizers were modified by the introduction of an additional baffle, as shown in figure 2, and used for atomization of washed particulate suspensions in distilled water. Under these conditions, the particle size is uniform and is determined by that of the suspended particles. Satisfactory uniformity of the aerosol has been shown by study of impinged samples by light-field, dark-field, and phase-difference microscopy. Less than 5 per cent of the particles appeared as doublets; no clumps, and only traces of very minute foreign particles were observed. Breakage of dead bacteria used in impingement studies is probably insignificant at jet velocities less than 150 m. per second. These points deserve emphasis.

In constructing the atomizing and diluting system shown diagrammatically in figure 3, we were aided by suggestions from Gucker and his group. As in figure 3, filtered air at constant pressure maintained by two Mason-Neilan regulating valves in series enters the system at the left, passing first through a copper coil, then through distilled water in a Pyrex gas-washing bottle, and subsequently through a coiled trap (not shown) to the atomizer. These parts were immersed in a water bath at 28°C. in order nearly to saturate the atomizing air. Other parts of the dilution system were placed in a bath at 25°C. The

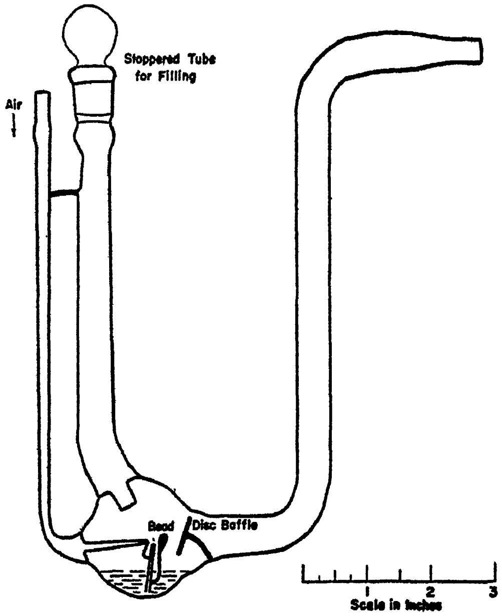


Fig. 2. Modified Vaponefrin nebulizer

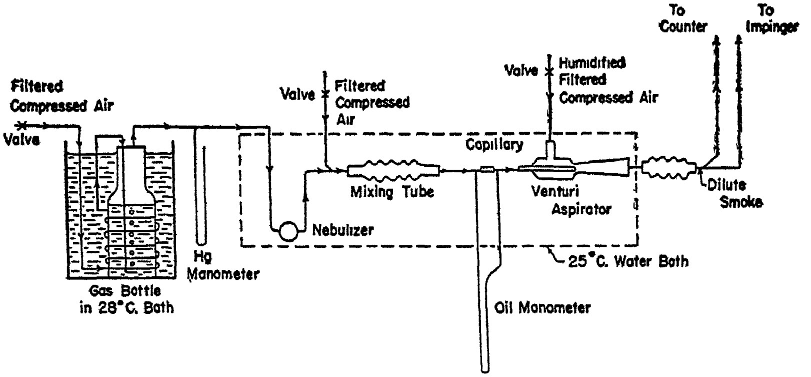


Fig. 3. System for aerosol generation and air dilution. Room temperature = 25°C.

atomizer mist was at once diluted with a measured volume of filtered air at known humidity, and passed through a mixing chamber consisting of a series of bulbs about 2 cm. in diameter. The mist evaporated almost immediately, and a small fraction of the aerosol passed through a capillary flowmeter into a venturi tube where it was again diluted with a known volume of air. At the end, simultaneous sampling with the photoelectronic instrument and impingers or other collecting devices was possible. Not shown on the diagram are outlets for the collection and disposal of excess aerosol. The excess eventually passed into a common duct in which it could be sterilized by ultraviolet radiation, or heat, or both.

The flow of diluting air was measured with calibrated capillary flowmeters. Dilution factors could thus be calculated from observed flow rates.

The humidity of the aerosol was manually controlled by mixing saturated and dry, or partly dried, air. It was measured with an Alnor Dewpointer. Since the dilution ratio in the final stage is 40:1, measurement of the humidity of the diluting air instead of the aerosol should be sufficiently accurate and safer when virulent organisms are used.

With this system we have prepared aerosols, composed predominantly (95 per cent upwards) of single bacterial particles at a concentration of 100–5000 particles or $5\text{–}250 \times 10^{-11}$ g. per liter. The relative humidity of the aerosols studied was generally about 62 per cent.

Counter operation: Since the photomultiplier tube had not been successfully applied in the Gucker instrument when these studies began, a Thalofide cell was used throughout. The procedure described by Gucker (38) was somewhat modified.

(*1*) Air flow through the smoke cell was half that given by Gucker. Smoke was sampled at 0.5 l. per minute; sheath air flowed at 1.5 l. per minute.

(*2*) The 6–8 v. tungsten filament bulb was supplied from four-cell storage batteries instead of three, producing a voltage drop up to 7.0 v. across the filament and markedly increasing light intensity. Although the lamps burned out after 4–8 hr., they could readily be replaced.

(*3*) A scale of sixty-four scaler was interposed between the phototube amplifier and the thyratron tube.

C. General

Those components of the apparatus in which the pressure was likely to exceed that of the atmosphere, including all of the atomizing and diluting system, were placed in a separate room which was maintained at slightly reduced pressure by ventilation with an exhaust fan. The effluent could be sterilized by heat or ultraviolet radiation. Leakage into the outer room was thereby avoided.

In the outer rooms were the electrical apparatus, vacuum pumps, smoke cell, and sampling outlet. Aerosols contained in parts of the smoke line passing into the outer room were generally at reduced pressure so that accidental leaks would be inward. Openings could be protected by efficient filters.

D. Preparation of bacterial suspensions

The preparation of bacterial suspensions is briefly described in appendix A.

E. Results

In figure 4 the photoelectronic count of *S. marcescens* aerosols appears as ordinate, and time as abscissa. The photoelectronic instrument rapidly follows the concentration effect observed with reflux atomizers. The solid lines are calculated in terms of atomizer delivery at constant air pressure and volume, temperature, and water content of the expanded atomizing air, by means of a simple equation of the form

$$C_t = C_0 \frac{V_0 - at}{V_0 - bt}$$

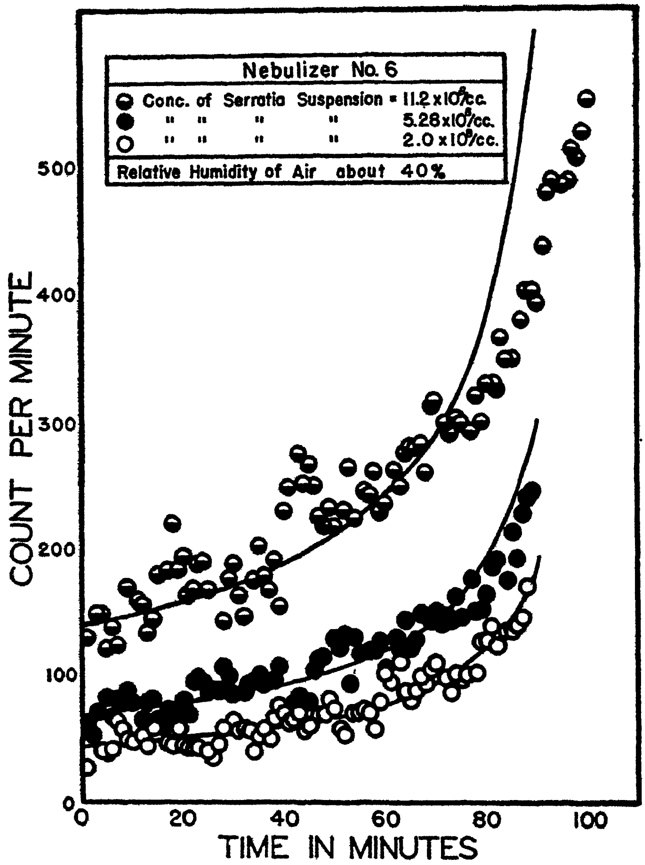


FIG. 4. Instantaneous direct photoelectronic measurement showing increasing aerosol concentration.

where C_t is the count at time t in minutes, C_0 is the initial count, b is the observed weight loss per minute at observed temperature, humidity, and air pressure, a = (b minus the evaporative loss sufficient to saturate the initial mist).

The values a and b are characteristic for each atomizer, and their systematic variation with air pressure, temperature, and humidity could be computed. They may, however, be obtained by atomizer calibration. The computations suggest that constant aerosol concentration might be attained by adding water to the liquid suspension at a rate equal to that of evaporative loss.

The parallelism of the concentrations of the aerosol and of the bacterial suspension in the atomizer has been shown by comparing direct bacterial counts made at the beginning and end of an experiment[3]. These ratios, compared with those of the corresponding photoelectronic counts, are, within the experimental error, unity. They do not show how completely the photoelectronic instrument enumerates bacterial particles. Other methods are necessary for this determination.

With *B. globigii* suspensions, the course of events is similar, but sometimes obscured by spontaneous clumping in the suspension which partially blocks the atomizer. These concentration phenomena are surely significant when atomizers of this type are used to administer drugs as aerosols.

II. COMPARISON OF THE PHOTOELECTRONIC INSTRUMENT WITH AN INDEPENDENT PHYSICAL METHOD

A. Introduction

Comparison with an independent physical method was desirable in order to find out whether photoelectronic counts were complete, and to determine the smallest particle size at which satisfactory counts were possible. For this purpose, comparison with biological methods seemed inappropriate, since possible changes in the viability of suspended organisms would introduce uncertainty. In order to avoid errors due to changes in concentration, it was necessary to compare samples taken simultaneously.

Inasmuch as the Gucker instrument, at least in its original form, had a frequency response limited to 1000 counts a minute (38), particulate concentrations could not exceed 1000 per liter at a flow rate of 1 l. a minute, or 2000 per liter at half this rate. Consequently many existing methods, most of them reviewed by Drinker and Hatch (21) and by Heymann (44), were not suitable for making comparisons. The methods described by them utilize the following principles for collection of particles from the air: (*1*) thermal precipitation, (*2*) filtration, (*3*) electrical precipitation, and (*4*) impingement. (*5*) In addition, we seriously considered a direct count of particles containing radioactive materials by means of a Geiger counter. (*6*) Use of an electrostatic counter devised by Guyton (41) appeared unsuitable, since it is not sufficiently sensitive (39).

In the first four methods, estimation of concentration is derived either from microscopic observation, weighing, or analysis; in the last two, directly from air flow and count of electrical impulses.

(*1*) For our purposes *thermal precipitation* (2, 35, 91, 96) appeared unsuitable, since the sampling rate of about 6 cc. per minute would introduce large sampling errors at low concentrations unless the sampling period were prolonged. In that case, errors attributable to changing aerosol concentrations were certain unless control of concentration was precise. (See also Rosenblatt and La Mer (77a).)

(*2*) Filtration also appeared not to provide a satisfactory approach, since quantitative recovery of particles from either soluble or insoluble filters did not seem assured (77).

---

[3] Ferry, Farr, and Hartman: Unpublished data.

(*3*) Electrical precipitation (6, 22, 73, 82), at least without considerable modification, appeared to spread a deposit over so wide an area that accurate direct counts were likely to be difficult at the low aerosol concentrations studied. Washing off the deposit, followed by concentration and subsequent enumeration, seemed a precarious approach.

(*4*) Impaction or impingement on a surface seemed promising. There are numerous variants such as:

(a) Simple settling: This is slow and ineffective with particles as small as 1 μ or less (33).

(b) Centrifugation: Wells' centrifuge is intended primarily for cultural studies. With it, collection efficiency appears uncertain (14).

(c) Bubblers }
(d) Washing } Neither of these methods is adapted for the study of dilute aerosols, composed of dead bacteria, since particles cannot be readily counted without concentration and transfer (36, 37). It is difficult to be certain that this procedure is rigorously quantitative.

(e) Impingers: The studies of Owens (67) and, among others, the more recent studies of May (61) and Sonkin (79) suggested that at appropriate jet velocities particles somewhat smaller than 1 μ in diameter, and with a density near unity, might be nearly quantitatively recovered and counted. This method was therefore used for appraising the photoelectronic counter.

(5) We seriously considered the use of a Geiger counter to count particles containing radioactive material. Specious counts due to the presence of minute radioactive particles outside the bacteria were possible, and proof of their absence seemed difficult. It was not attempted.

B. Methods

Impinger details: After preliminary experiments had shown that neither a cascade impactor (61, 79) nor a Bausch and Lomb dust counter (40) produced deposits sufficiently symmetrical or well defined to make accurate particulate count possible, the impinger shown in figure 5 was made. In its present form, it resembles Owens' jet impinger (67) and consists of a simple slit jet which impinges on a glass disc held at a fixed distance from the orifice.

It is composed of two main parts, one containing the jet and slide holder, the other a funnel-shaped piece through which the aerosol leaves the collector. The parts can be tightly closed by a large nut which fits over the outer threaded surface of the funnel-shaped piece. Consequently small metal particles, produced by tightening the threads, cannot be caught in the air stream and deposited in a second impinger. Sharp corners have been eliminated to minimize deposition except on the collecting slide. The design of the collecting chamber is such that air flow divides symmetrically on either side of the jet orifice. Stainless-steel construction permits frequent sterilization.

Two types of jet, both shown in figure 6, were tried. The tapered jet shows less apparent constriction and probably induces less turbulent flow. None the

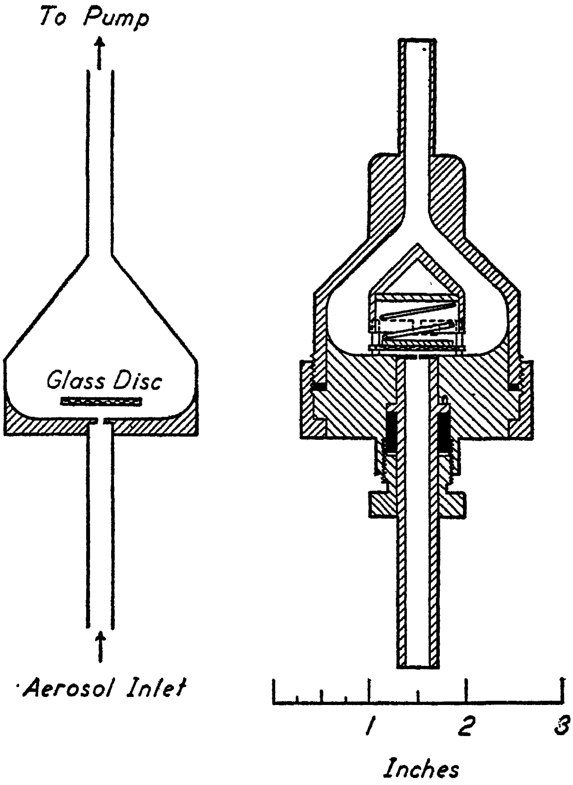


FIG. 5. Stainless-steel impinger: left, schematic drawing; right, scale drawing

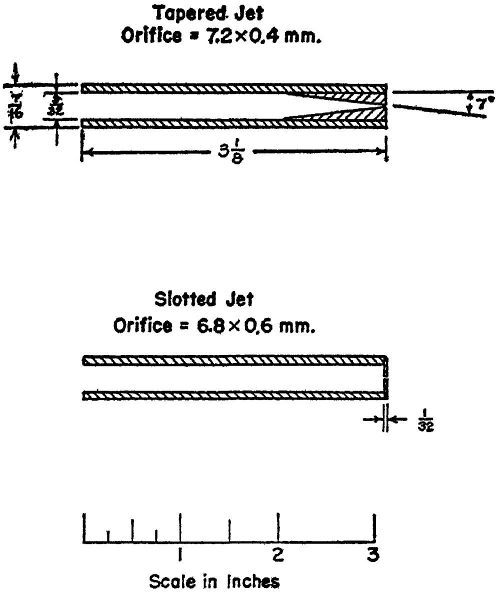


FIG. 6. Impinger jets

less, the simple rectangular slit 6.8 x 0.6 mm., with carefully polished edges, resulted in about 1.5 times better recovery.

Calibration was in terms of air flow, measured in liters per minute with a standard rotameter accurate to about 3 per cent, and pressure, in millimeters of mercury, at room temperature 25°C. ± 0.5°. The effective cross-section and the apparent constriction factor were calculated from the observed air flow at critical pressure drop. The latter turned out to be about 0.8, a reasonable value. Average linear jet velocities were, in turn, calculated from the effective cross-section, air flow, and pressure–volume relations.

Although the jets were intended to be alike, one was slightly larger than the other. The larger orifice was placed second when two impingers were used in series, with the result that calculated jet velocities differed by not more than 3 per cent at air flow up to 21 l. per minute. At 25 l. per minute the difference was about 12 per cent.

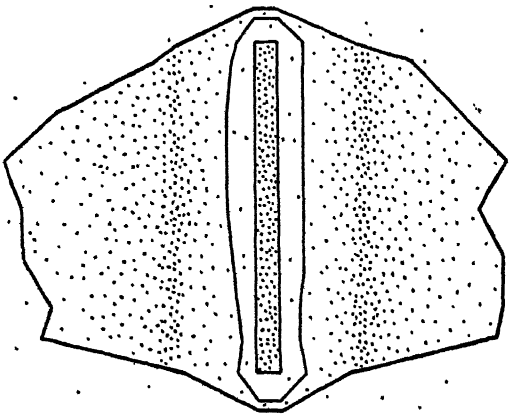


FIG. 7. Diagrammatic representation of a typical impinger deposit. Jet deposit and halo pattern for slotted jet.

Air flow through the impingers was actuated by a mechanical vacuum pump having a capacity of 12 cu. ft. per minute. Flow was regulated by calibrated, critical orifices placed between the impingers and a filter which prevented escape of live organisms to air—a vital precaution with pathogenic bacteria. With single impingers, flow rates up to 35 l. per minute, and consequent jet velocities up to 345 m. per second were possible. Two impingers used in series reduced the maximum air flow to 29.7 l. per minute.

C. *Collection and count of impinger samples*

Collection of particles: Particles were collected on glass discs (23 x 1.5 mm.). These were scrupulously clean and free from scratches. Metal spacers permitted separation from the orifice by 0.25–3.5 mm. In order to decrease the chance that particles would strike an area already filled by an organism and carom, exposure in the air stream was only long enough to collect about 5×10^4 particles.

Under these conditions and at optimum flow rates, the impinged deposit takes the form shown in figure 7. A central primary deposit, almost exactly the size of the orifice, is immediately surrounded by a clear space containing a few scattered organisms. This is, in turn, surrounded by a halo of impinged particles, which is sharply defined. Only occasional organisms appear beyond its limits. The density of the secondary deposit or halo is not uniform, as is shown in figure 8, in which the average density of individual strips is plotted as ordinate and the distance from the center of the deposit as abscissa. A smooth curve has been drawn through points obtained from two experiments. This, with the observation that particles were not impinged on cover slips fixed to the "roof" of the collector, suggests that deposition is essentially restricted to the area outlined in figure 7. Our confidence that estimated aerosol concentration based on impinger collection approximates an actual value rests primarily on this evidence and appears well founded.

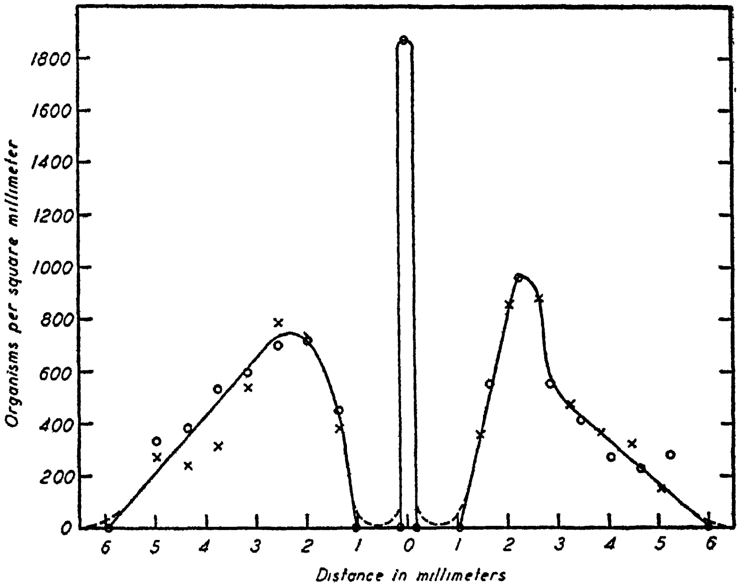


FIG. 8. Intensity of deposit in relation to lateral distance from orifice

Enumeration of impinger samples: The number of particles deposited was estimated from an average deposit density per square millimeter (derived from the count of representative strips at right angles to the jet) and the area of the deposit, determined by measurement and planimetry of a scale drawing thereof. The area times the density, divided by the total air flow, gave the number of particles collected per liter of air, a number directly comparable to that obtained photoelectronically.

Representative strips were defined by a slit eyepiece, and particles counted by means of dark-field illumination with a high-power dry lens. The length of each strip was measured with the aid of a mechanical stage; its width was determined by calibration of the eyepiece slit with a stage micrometer. The density of the particulate population per square millimeter was separately determined for each strip, and a mean density derived. This value multiplied by the area of the deposit yielded the total number deposited. From it, the particulate concentration per liter was obtained by dividing by the total air flow.

Particulate population, not only of the primary deposit but also of the right and left portions of the halo, was estimated separately. The ratios of the population of the jet deposit to that of the halo were not sufficiently constant to justify estimates based solely on counts of the former. But when two impingers were used in series, there was no statistically significant difference between these ratios as computed for first and second impingers. We believe that this supports the microscopic observation that particle size is approximately uniform.

Experiments did not indicate that total counts were markedly increased by coating slides with an alkyd resin adhesive. The density of the primary deposit appeared, however, to be slightly increased at the expense of that of the halo. Since coating tended to interfere with dark-field microscopy, it was abandoned.

D. Results

Single impingers: Results of a number of experiments performed with *B. globigii* aerosols using single impingers are graphically shown in figure 9, in which calculated jet velocity appears as abscissa and the counter/impinger count ratio as ordinate. The full line represents results obtained with slit jets; the broken line those derived with tapered jets. In the former case, maximum collection efficiency lies at velocities between 90 and 120 m. per second, and appears to be slightly more than 50 per cent. At velocities greater than 150 m. per second, efficiency rapidly diminishes.

The observations represented in figures 7, 8, and 9 suggest that impingement should be largely explicable in terms of microaerodynamics within the impinger. Comparison of the efficiency of tapered and slit jets suggests that turbulence favors impingement. Although the geometry of the deposit and the increased size of both clear space and halo at higher jet velocities are striking, we can only offer a qualitative explanation. This phenomenon, the less effective impingement of solid as compared with liquid droplets observed by Landahl (51) and May (61), and the reduction in collection efficiency when jet velocities exceed optimal values, can all be qualitatively explained by the assumption that these particles are elastic. Some of those striking the surface under the jet would bounce; of these, a fraction could be returned to the halo and adhere; some would be carried by the air stream and, as its velocity increased, would clear the slide. Many would be deposited elsewhere. The marked diminution of deposition observed at jet velocities above 150 m. per second might also be due to shattering of bacteria into very fine particles, not microscopically visible. But the relative importance of aerodynamic factors, elasticity of the particles, and shattering cannot yet be quantitatively evaluated.

Two impingers in series: Although impingers were occasionally used singly, rather more consistent data could be obtained when two collectors were placed in series. Data representative of an experiment with two impingers in series appear in table 1.

The experimental conditions are given at the top of the table. Flow rates in liters per minute, determined by means of critical orifices, appear in the second column. Calculated jet velocities preceded by the numbers (1) and (2) refer to velocities at the first and second orifices, respectively. Spacer thickness, given in millimeters, indicates the distance between collecting slide and orifice. Under the heading "Concentration per liter, impinger," sets of three values generally appear. These sets of values from top to bottom denote successively the concentration derived from the first impinger deposit alone, from the second deposit, and from their sum. Figures in parentheses indicate an estimate with patently ineffective collection, and a single value is given where an impinger was used singly. "Concentration per liter, counter" refers to photoelectronic enumeration. The ratios in the final column are usually paired; the upper value refers to the ratio between the electronic count and the number collected by the first impinger; the lower expresses the relation between counter value and the num-

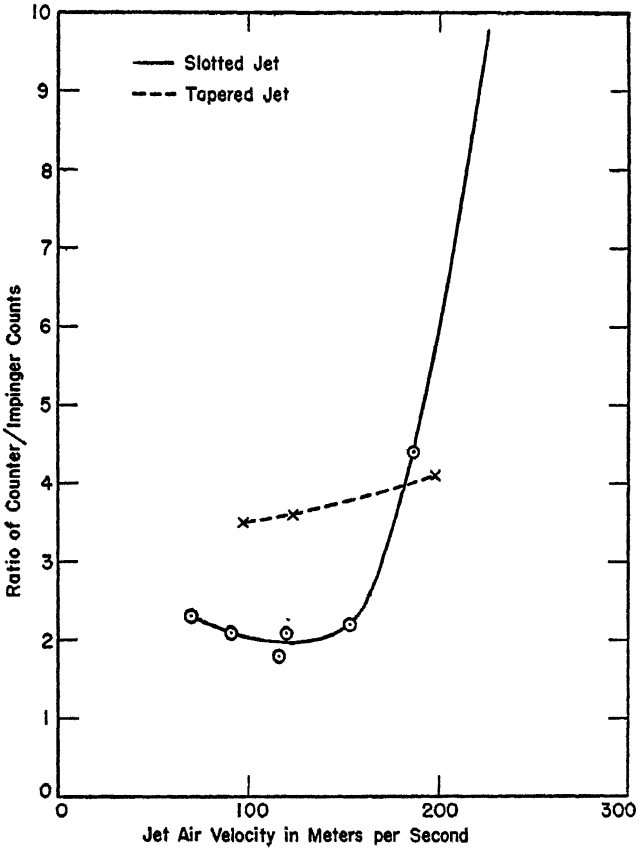


FIG. 9. Comparison of collection by single impingers with direct photoelectric measurement

TABLE 1

Comparison of impinger and photoelectronic counts

February 13, 1948

Conditions

Barometer (cor.), 765.6 mm. Temperature: room, 24.5°C.; bath, 25°C.	Dewpoint of atomizer air, 16.8°C.; aerosol, 16.2°C.	Organism, *B. globigii* spores Concentration of suspension, 1.6 × 10^8 cc.

RUN	FLOW RATE	CALCULATED JET VELOCITY	SPACER THICKNESS	CONCENTRATION PER LITER		RATIO COUNTER/ IMPINGER
				Impinger	Counter	
	l./min.	m./sec.	mm.			
1	16.8	(1) 91	0.5	572		2.4
		(2) 92		527	1360	
				1099		1.2
2	16.8	(1) 91	0.5	707		2.0
		(2) 92		367	1418	
				1074		1.3
3	20.8	(1) 118	0.5	739		2.0
		(2) 122		370	1462	
				1109		1.3
4	20.8	(1) 118	0.5	626		2.3
		(2) 122		334	1464	
				960		1.5
5	20.8	(1) 118	1.0	570		2.5
		(2) 122		304	1398	
				874		1.6
6	20.8	(1) 118	1.0	605		2.3
		(2) 122		331	1398	
				936		1.5
7	25.2	(1) 149	1.0	478	1402	2.9
		(2) 168		(60)		
8	25.2	(1) 149	1.0	549	1320	2.4
		(2) 168		(69)		
9	16.8	(1) 91	2.0	477	1326	2.8
		(2) 92		245		
				722		1.8
10	16.8	(1) 91	2.0	474	1352	2.9

ber collected by two impingers in series. Where no second value is given, collection by the second impinger was either negligible or the deposit faulty.

Collection by the first impingers in replicate runs may differ by as much as 25 per cent, presumably owing to unpredictable changes in the air streams within the impinger resulting from minute changes in the position of the collecting discs. The number of particles collected by two impingers in series in two replicate runs appears to differ less markedly. Indeed, the error is consistent with the expected sampling error ±35, when the count is about 1000.

In table 2 appear results derived from all experiments with *S. marcescens* as well as *B. globigii*, in which the light bulb was supplied by an 8-v. storage battery.

Ratios of concentrations determined photoelectronically and by the impinger method are given, including values for single impingers, for two in series, and a calculated value for 100 per cent collection by the impinger method. At the extreme right of the table, the values of the ratio indicate progressive diminution of collection efficiency as the distance between slide and orifice increases. This confirms the observations of Bourdillon and others (10).

TABLE 2

Comparison of aerosol concentrations determined photoelectronically and by impingement

SPACER THICKNESS (MM.)...		0.5			1.0	2.0	3.5
APPARENT VELOCITY	TYPE OF JET	RATIOS: $\frac{\text{PHOTOELECTRIC CONCENTRATION}}{\text{IMPINGER CONCENTRATION}}$					
		One impinger	Two impingers in series	$\frac{\text{Photoelectronic}}{\text{Calculated}}$	Single impinger		
(A) *B. globigii*							
m./sec.							
113	Slit	1.8				2.80	
118	Slit	2.1±0.23	1.4±0.14	1.1±0.10	2.40	2.70	(50)
149	Slit	2.2			2.70		
123	Taper	3.6	2.1	1.5			
(B) *S. marcescens*							
118	Slit	2.9	2.0				
149*	Slit	2.2	1.4	1.0±0.1			
233†	Slit	2.8					

* Velocity at second jet 169 m. per second.

† Velocity at second jet sonic.

The calculated value rests largely on the theoretical conclusions reached independently by May (61) and Sell (78) and applied by La Mer and Hochberg (49) and Johnstone, Winsche and Smith (46) to the deposition of larger particles from slower air streams. May's data (61, fig. 7, p. 192) suggest that one may write impingement efficiency, I, as

$$I = \frac{\text{Number of particles impacted per liter of aerosol}}{\text{Concentration, particles per liter}} = \frac{N_i}{C_a} = f\left(\frac{\rho V D^2}{\eta l}\right)$$

where P denotes density, V is velocity in centimeters or meters per second, D is the particulate diameter in centimeters, η is the viscosity in poises, and l is a linear dimension to which May and Sell attach different physical significance. The experimental conditions, including humidity, are such that ρ, η, and l are constant, so that we may write,

$$I = f(KVD^2)$$

Each bacterial species is approximately uniform in size, and although not spherical, we take D^2 as a characteristic constant dimension related to surface. May's data, moreover, suggest that I is nearly linear in KVD^2 over the range 25–75 per cent, or in KV, where D is constant. Consequently, since aerodynamic conditions are similar in both impingers, we write:

$$N_1 = C_aKV_1 = N_0I$$
$$N_2 = (1 - C_aI)IV_2/V_1$$

and solving

$$I = 1 - \frac{N_2}{N_1} \times \frac{V_1}{V_2}$$

The meaning of the symbols is unchanged, except that the subscripts refer to the first and second impingers, respectively.

For *B. globigii* $I = 0.51$ and $\sigma = \pm 0.07$, so that two impingers in series should remove 76 per cent of the suspended particles. Total concentrations may readily be calculated. Values for the mean value of the ratio

$$\frac{\text{Photoelectronic concentration}}{\text{Calculated concentration}} = \frac{C_p}{C_i}$$

appear in the fifth column of table 2. Since the impinger method appears likely to provide slightly low estimates of concentration, the value of the ratio

$$C_p/C_i = 1.1, \qquad \sigma = 0.1$$

is very satisfactory. We believe that it implies that photoelectronic measurement of aerosol concentration is accurate when particle diameter is somewhat more than 0.7 μ.

With *S. marcescens*, similar considerations lead to values of $K = 0.47$, 75 per cent collection efficiency for two impingers in series, and

$$C_p/C_i = 1.0, \qquad \sigma = \pm 0.1$$

These values are, however, less certain than those for *B. globigii*, since the jet velocities of 150 and 170 m. per second, respectively, for the first and second jets may differ sufficiently to invalidate the assumption that collection is linear in the velocity. Moreover, the estimates of velocity are less accurate, since pressure head, on which they depend, is changing rapidly in this range. None the less, comparison of the calculated values for *B. globigii* and *S. marcescens* suggests that photoelectronic count is about 90 per cent complete with the smaller organism.

The values for impingement efficiency are lower than estimates derived from the collection of similar droplets of methylene blue and glycerol by means of a cascade impactor (79). This is to be expected, since collection of liquid particles is more complete (51, 61). Moreover, these liquid droplets are more dense than the bacterial particles on which our estimates rest.

Recent preliminary observations suggest that the counter may be more readily calibrated. If light intensity is increased by gradually increasing filament voltage from 2.0 to 7.0 v., the counting rate increases rapidly at first, later approaching a limiting value. If the count per minute is indicated as ordinate and voltage as abscissa, the resulting curve is sigmoid. With *B. globigii* the limiting value appears to be reached when filament voltage is not over 5.5 v. With *S. marcescens* the count appears to be about 90 per cent complete at 6.5 v., and is not quite complete at 7.5 v.

III. THE DETERMINATION OF THE CONCENTRATION OF BACTERIAL AEROSOLS BY THE METHOD OF COLLECTION AND CULTURE, COMPARED WITH DIRECT PHOTOELECTRONIC COUNT

A. Introduction

The determination of the concentration of living bacteria in air has depended on their collection, subsequent culture, and enumeration of the colonies appearing in plate cultures. Although biological material can be collected by methods used with inanimate matter, care is necessary to avoid impairment of viability. The methods generally used depend either on (*1*) filtration (63), (*2*) electrostatic sampling (9, 56), (*3*) bubbling through a liquid, (*4*) impingement (10, 77), (*5*) centrifugation (92), or (*6*) some combination of these methods (53, 56, 64). Filtration appears unsatisfactory, since recovery of viable vegetative organisms is poor (77). Electrostatic sampling (9, 56) appears promising, but has not, as far as we know, been carefully compared with other methods. Application of other methods is such that collection is either directly on a solid culture medium (10, 56) or into a liquid, from which aliquots are taken for culture (64, 74, 77), mixed with molten agar, and poured into plates. In either case, colonies are counted after an appropriate period of incubation at a temperature suitable for the organism studied. When collection is initially in a liquid medium, the concentration of viable particles in the aerosol, but not necessarily of viable organisms, is derived as follows:

$$\text{Concentration, particles per liter} = \frac{\text{volume of sample (cc.)}}{\text{volume of aliquot (cc.)}} \times \frac{\text{number of colonies in culture}}{\text{air flow (l./min.)} \times \text{period of sampling (min.)}}$$

If collection is directly on a solid medium, the term including volumes of sample and of aliquot is unnecessary. The relative collection efficiency of a number of typical devices has been compared by du Buy, Hollaender, and Lackey (14). One of these is an atomizer bubbler, devised by Moulton, Lemon, and Puck (64), in which the aerosol atomizes some of the collecting fluid. The mist generated is collected by bubbling through fluid in a second reservoir. Although bacterial concentrations measured with it are generally higher than those obtained with other devices, this is attributed to fracture of clumps of organisms. Results with the air centrifuge were not consistent. Moreover, both Lemon (53) and Rosebury (77) have found that impinger-bubblers collect nearly as efficiently as the Moulton device and are much more convenient.

Provided suitable precautions, described in modern bacteriological texts, are observed, culture counts are accurate to about 5 per cent. The error depends on the population of the sample taken.

B. Comparison of the bacterial concentration derived by the method of collection and culture with direct estimation by photoelectronic count

It seemed that the method of collection and culture might be quantitatively appraised by comparison with the photoelectronic method, which directly measures the total particulate concentration of bacterial aerosols. Provided the bacterial aerosols were pure, and provided further that the ratio of viable organisms to the total number present, commonly termed the viability coefficient, were known, absolute collection efficiency could be estimated. Or, alternatively, it would at least be possible to follow changes in viability by comparison of the photoelectronic and biological counts. The methods used and the results attained are next discussed.

C. Methods

1. Photoelectronic counts were made according to procedures already indicated.

2. The determination of viability ratios is essential for this proposed comparison. The methods used are described in appendix B.

Collection and culture of airborne organisms was accomplished by the method described by Rosebury (77), with impinger-bubblers operating as critical orifices at an air flow between 2 and 2.5 l. per minute.

The orifice was immersed about 4 mm. below the liquid level and 5 mm. above the bottom of thick-walled Erlenmeyer flasks into which they were inserted. Although 25 cc. of collecting fluid and four drops of olive oil were placed in each flask before sterilization, fluid losses amounting to about 2 cc. occurred during this process and the course of a run. Losses and final liquid volumes were determined by weight differences. After collection, the concentration of organisms per liter of aerosol was estimated by pour-plate colony counts.

D. Results

Representative data appear in table 3, and include zero values for the controls. The captions are self-explanatory. The data presented serve to illustrate (*1*) the consistency of data obtained, (*2*) the fact that clumps are broken up during atomization, and (*3*) especially that no significant changes either in apparent or true viability coefficients occur during an experiment, despite repeated atomization in the refluxing atomizer used. A summary comparing estimates of total

TABLE 3

The viability of bacterial suspensions as affected by atomization
April 9, 1948; organism, *S. marcescens*

A. SUSPENSION CHARACTERISTICS

	Before experiment	After experiment
Particulate concentration in atomizer $\times 10^{-8}$	1.0	1.4
Concentration of viable organisms in atomizer $\times 10^{-8}$	0.6	0.9
Per cent organisms as doublets	60	46
Apparent viability (per cent)	55	64
"True" viability (per cent)	60	68

The collection of living organisms from bacterial aerosols

B. AEROSOL CHARACTERISTICS
TIME OF SAMPLING, 2 MIN.; FINAL VOLUME OF COLLECTING LIQUID, 23 CC.

Run	Flow rate	Volume of aliquot: 1.0 cc.	Volume of aliquot: 0.5 cc.	Volume of aliquot: 0.2 cc.	Concentration of organisms per liter
		Number of colonies per plate			
	l./min.				
I	2.27 (control)	0	0	0	0
II	2.27	54	27	9	370
III	2.28	54	25	12	392
IV	2.28	49	25	10	363

TABLE 4

Comparison of particulate concentrations of bacterial aerosols estimated by biological methods and by direct photoelectronic count

S. marcescens					*B. globigii*				
DATE	RUN	CONCENTRATION: PARTICLES PER LITER		APPARENT COLLECTION EFFICIENCY	DATE	RUN	CONCENTRATION: PARTICLES PER LITER		APPARENT COLLECTION EFFICIENCY
		Biological	Direct				Biological	Direct	
				per cent					*per cent*
April 1, 1948	II	535	1406	38	April 13, 1948	II	775	2098	37
	III	440	1654	27		III	844	2020	42
	IV	447	1728	26		IV	982	2150	46
April 9, 1948	II	370	1425	26	April 14, 1948	II	911	2106	43
	III	392	1518	26		III	934	2064	45
	IV	363	1555	23		IV	896	1964	46
Mean				28	Mean				43

bacterial or particulate concentration of aerosols derived by the method of collection and culture with direct photoelectronic determinations appears in table 4.

Since no significant change in viability appeared to result from atomization, the mean values determined for the aqueous bacterial suspensions in the atomizer have been assumed to apply to organisms in the aerosol. These values are taken as 0.7 for *S. marcescens* and 1.0 for *B. globigii* spores. The total bacterial concentration, based on the biological method of collection and culture, can be obtained by dividing the viable count by the viability coefficient. With the counter, particulate concentration is directly obtained. Both values appear under the captions "Concentration: particles per liter." The subheadings "Biological" and "Direct" distinguish them. Since the evidence discussed in Section II (page 406) suggests that the photoelectronic counter completely enumerates *B. globigii* and *S. marcescens* to the extent of 90 per cent, we have termed the ratio of biological concentration to direct concentration, "apparent collection efficiency."

The values are consistent, though smaller than those calculated from the impingement of dead organisms with another device. This is not surprising, since aerodynamic conditions are different, the jet velocities (345 m. per second) are greater than optimal, and the mean distance from jet to impinging surface is excessive. Some bacterial breakage may occur as the result of impingement at sonic velocity. Its extent probably varies with bacterial species, but has not yet been determined.

The efficiency is, however, less than that estimated by Rosebury (77) for these organisms, using similar impingers. This discrepancy is not explicable in terms of excessively high particulate counts resulting from contamination by non-bacterial particles, nor is it likely to be due to impaired viability of atomized organisms. Rosebury's experimental arrangements were, however, so different that the discrepancy may well be attributed to differences in density and particle size.

DISCUSSION

The simple theoretical considerations of May (61) and Sell (78) permit comparison of impingement efficiencies derived from physical and biological measurement. When impingement efficiencies are identical, the jet velocities should be inversely proportional to the squares of representative diameters; at identical velocities, the ratio of squared diameters should equal that of the efficiencies, since $I = KVD^2$.

In table 5 we have collected appropriate values for jet velocity, V, diameter, D, and efficiency, I.

For example, the method of microscopic count shows that collection efficiency is 45 per cent with *S. marcescens* at 150 m. per second; with *B. globigii*, we estimate from table 2 and figure 9 that 45 per cent efficiency is attained at a velocity of about 88 m. per second. The velocity ratio is about 1.7.

We should expect impingement to be proportional to the surface of that part of the particle that comes in contact with the collecting medium. In the case of spheres, this is proportional to D^2, where D is the diameter. But the bacterial particles, although symmetrical, are not spherical. For these ellipsoidal particles, comparable values may be obtained if we calculate the diameter of the sphere of equivalent mass. This appears in the table under the heading "diameter." The assumed mean values for particulate diameters may, of course, differ somewhat from the true values. Tentatively basing comparison on the assumed dimensions of *B. globigii*, we have computed a diameter for a sphere equivalent to *S. marcescens* which will exactly accord with theory. This turns out to be 0.7 μ instead of 0.66 μ calculated from axes 0.6 μ and 0.8 μ.

But in a streaming medium particles tend to orient themselves so that the long axis is parallel to the line of flow. One would then expect the impinging surface and impingement to be proportional to the squares of the short axes of an ellipsoidal particle. Values based on this assumption appear under the heading "Short axis of ellipsoid." Agreement with the observed value is somewhat closer. It has, however, been shown that the counter values are 10 per cent low in the case of *S. marcescens*. The observed value for collection efficiency should therefore be correspondingly reduced. The corresponding ratio turns out to be 1.7, in excellent agreement with theory.

TABLE 5

Theoretical appraisal of impingement data

	MEAN DIAMETER D IN MICRA				PHYSICAL METHOD		BIOLOGICAL METHOD		
	Sphere of equivalent mass		Short axis of ellipsoid		I	V	I		V
							Uncorrected	Corrected	
					per cent	*m./sec.*	*per cent*	*per cent*	*m./sec.*
Organism:	D	D^2	D	D^2					
B. globigii	0.92	0.85	0.8	0.64	45	88	43	43	345
S. marcescens	0.66	0.44	0.6	0.36	45	150	28	25	345
Ratio $R = \frac{B.\ globigii}{S.\ marcescens}$	$R = 1.9$		$R = 1.8$		$1/R = 1.7$		$R = 1.5$	$R = 1.7$	

Although this hypothesis seems attractive, our data are too scanty for definitive proof. Further study of the impingement, both of spherical particles, such as the sulfur particles described by La Mer (48), and of more elongated bacteria, should provide information that has theoretical and practical significance.

The theory of impingement is certainly incomplete in other respects. Indeed, the significance of the dimension, l, appearing in the denominator appears in doubt. May (61) relates it to slit width; Sell (78) to a dimension of the object on which impaction takes place. The former explanation appears unsatisfactory, since this dimension is directly related to jet velocity, and with the impinger used by us, Sell's explanation appears irrelevant. That it is not directly related to slit-to-slide distance has been indicated in table 2.

Further information is necessary concerning the influence of physical properties of the particle and the collecting surface such as elasticity, number and orientation of charges, and "tackiness." The elasticity and friability of bacteria, both dead and alive, deserve further study.

Although the effect of long-range forces may well be more important, the effect of charge can be experimentally determined by coating slides with substances differing in this respect. Although particle size and also shape may, as suggested by Gucker and by Mueller (65), be determined by photoelectronic measurement of the intensity of light scattered at different angles, it should also be possible to estimate particle size with the present instrument from the relation between counting efficiency and intensity of the light source. Results must, however, be carefully interpreted, since Barnes and La Mer (7) have shown that the intensity of light scattered by particles is characterized by a function in which the exponent changes with the ratio of particle size to the wave length of the scattered light.

Empirical calibration may be easier by means of aerosols generated from the extraordinarily uniform sulfur sols prepared by La Mer (48). Such aerosols should also demonstrate the effect of changes in particulate refractive index on the intensity of scattered light. As the index is increased, it should be possible to count smaller particles accurately.

But even in its present form the photoelectronic instrument instantaneously provides information which can be obtained by other methods only after hours of meticulous and tedious observation. With improvements based on the use of more sensitive photocells and changes in the optical system, it should be possible to determine accurately the concentration of aerosols containing particles considerably smaller than 0.7 μ in diameter, the present lower limit.

And even with this limitation, it is a useful tool for the rapid estimation of the concentration of dilute bacterial aerosols. It should be useful for calibrating methods for the collection of particulate matter from aerosols, for the determination of changes in viability, and in quantitative laboratory investigation of airborne disease, especially those concerned with penetration and dosage.

In conclusion, it should be emphasized that the counter is at present a laboratory instrument, that enumeration of particles smaller than 0.7 μ is incomplete, and that it cannot be directly applied to determining the concentration of aerosols greater than 1000 particles per liter and therefore is not adapted for use in ambient air, which generally contains many particles. It is not sufficiently sensitive to detect and enumerate primary virus particles, some of which are only 0.01 μ in diameter. Biological fluids are, however, rarely as dilute as 0.1 per cent. The particulate size of infectious particles may often be largely conditioned by the concentration of the solute in the fluid in which they are suspended. One might expect particles about 1 μ in diameter to result from drops 10 μ in diameter if concentration of the solute is only 0.1 per cent, assuming that the solute has a density of 1.0. This fact may extend the utility of the instrument.

Moreover, improvements which will materially decrease the lower limit of size and increase the limiting concentration at which counts are complete seem theoretically and practically possible. With improvements, it should be possible not only to count suspended particles but also to measure their size and approximately to describe their size and shape.

SUMMARY

1. Since it has long been suspected and it is now definitely established that disease may be airborne, it is not surprising that medical scientists have studied and clarified the understanding of aerosols.

2. Certain bacterial species may be grown so that size and shape are nearly uniform. The study of bacterial aerosols exemplifying suspensions containing solid particles of low density should therefore be profitable on physical-chemical grounds as well as for the elucidation of the mechanism of transmission of airborne disease.

3. Effective study of aerosols must be preceded by the generation of pure suspensions of particles of uniform size and shape, and by accurate measurement of their concentrations.

4. The photoelectronic instrument used in these studies is very briefly described.

5. A method of generating dilute bacterial aerosols is briefly presented.

6. Comparison of aerosol concentrations, determined by photoelectronic and impingement methods, suggests that direct measurement by the photoelectronic device is accurate provided (*1*) the diameter of the bacterial particle somewhat exceeds 0.7 μ and (*2*) illumination is sufficiently intense.

7. The concentration of living organisms, determined by biological methods after collection with impinger-bubblers, is less than particulate concentration determined photoelectronically. Although this discrepancy arises in small part from the presence of non-viable organisms, it appears to be largely attributable to incomplete collection, presumably due to unfavorable conditions, possibly including breakage at the high impingement velocity used.

8. Impingement efficiency, as already shown by May and Sell, is proportional to particle density and surface. It is suggested that with ellipsoidal bacterial particles, presumably oriented in the air stream, the surface effective for impingement is determined by the magnitude of the shorter axis of the ellipsoid.

9. A quantitative theoretical description of impingement must also include the effect of other physical properties, such as elasticity and friability of the particle, the magnitude and distribution of charge both on the particles and on the collecting surface, as well as the microaerodynamics of the impinger.

10. Advantages and limitations of the photoelectronic instrument are indicated. Improvements should decrease the particle size at which accurate measurements are possible.

REFERENCES

(1) *Aerobiology,* Publication No. 17, Forest R. Moulton, *Editor,* American Association for the Advancement of Science, Washington, D. C. Science Press, Lancaster, Pennsylvania (1942).
(2) Aitken, J.: *Collected Papers.* Cambridge University Press, London (1923).
(3) Andrewes, C. H., and Glover, R. E.: Brit. J. Exptl. Path. **22**, 91–97 (1941).
(4) Dautrebande, L., Highman, B., and Alford, W. C.: J. Ind. Hyg. Toxicol. **30**, 103–7 (1948).
(5) Baker, A. H. and Twort, C. C.: J. Hyg. **43**, 382–94 (1943).
(6) Barnes, E. C., and Penney, G. W.: J. Ind. Hyg. Toxicol. **18**, 167–72 (1936).
(7) Barnes, M. D. and La Mer, V. K.: J. Colloid. Sci. **1**, 79–91 (1946).

(8) *Bergey's Manual of Determinative Bacteriology*, R. S. Breed, E. G. D. Murray, and A. P. Hitchens, *Editors*, The Williams & Wilkins Company, Baltimore (1948).
(9) Berry, C. M.: U. S. Pub. Health Repts. **56**, 2044–51 (1941).
(10) Bourdillon, R. B., Lidwell, O. M. and Thomas, J. C.: J. Hyg. **41**, 197 (1941).
(11) Brown, C. E.: J. Ind. Hyg. **13**, 285–91 (1931).
(12) Brown, C. E.: J. Ind. Hyg. **13**, 293–313 (1931).
(13) du Buy, H., Arnold, F. A., and Olson, B. J.: U. S. Pub. Health Repts. **162**, 1391-1413 (1947).
(14) du Buy, H., Hollaender, A., and Lackey, Mary D.: U. S. Pub. Health Service, Suppl. Pub. Health Repts. No. **184**, 1–40, U S. Government Printing Office, Washington, D. C. (1945).
(15) Challinor, S. W.: J. Hyg. **43**, 16–54 (1943).
(16) Chope, H. D. and Smillie, W. G.: J. Ind. Hyg. Toxicol. **18**, 780–92 (1936).
(17) Dautrebande, L., Highman, B., and Alford, W. C.: J. Ind. Hyg. Toxicol. **30**, 103–7 (1948).
(18) Dautrebande, L., Alford, W. C., and Highman, B.: J. Ind. Hyg. Toxicol. **30**, 108–13 (1948).
(19) Debye, P.: J. Phys. Colloid. Chem. **51**, 18–32 (1947).
(20) Drinker, C. K.: J. Ind. Hyg. **3**, 295–303 (1922).
(21) Drinker, P., and Hatch, T.: *Industrial Dust.* McGraw-Hill Book Company, Inc., New York (1936).
(22) Drinker, P., Thomson, R. M., and Fitchet, S. M.: J. Ind. Hyg. **5**, 162–85 (1923).
(23) Duguid, J. P.. Edinburgh Med. J. **52**, 385–401 (1945).
(24) Duguid, J. P.: J. Hyg. **44**, 471–9 (1945–46).
(25) Dunklin, E. W., and Puck, T. T.: J. Exptl. Med. **87**, 87–101 (1948).
(26) Edward, D. G. F., Elford, W. J., and Laidlaw, P. P.: J. Hyg. **43**, 1–10 (1943).
(27) Edward, D. G. F., Lusk, D., and Bourdillon, R. B.: J. Hyg. **43**, 11–15 (1943).
(28) Edward, D. G. F., and Lidwell, O. M.: J. Hyg **43**, 196–200 (1943).
(29) Elford, W. J., and van den Ende, J.: J. Hyg. **44**, 1–15 (1945–46).
(30) Findeisen, W.: Arch. ges. Physiol. (Pflügers) **236**, 367–9 (1935).
(31) Finn, S. R., and Powell, E. O.: J. Hyg. **41**, 473–88 (1941).
(32) Flügge, C.: Z. Hyg. Infektionskrankh. **30**, 107–24 (1899).
(33) Gibbs, W. E.: *Clouds and Smokes.* The Blakiston Company, Philadelphia (1942).
(34) Graeser, J. B., and Rowe, A. H.: J. Allergy **6**, 415–20 (1935).
(35) Green, H. L., and Watson, H. H.: Medical Research Council, Special Report No. 199, London (1935).
(36) Greenburg, L., and Smith, G. N.: U. S. Bur. Mines Repts. Invest. No. **2392** (1922).
(37) Greenburg, L., and Bloomfield, J. J.: U. S. Pub. Health Repts. **47**, 654–75 (1932).
(38) Gucker, F. T., Jr., O'Konski, C. T., Pickard, H. B., and Pitts, J. N., Jr.: J. Am. Chem. Soc. **69**, 2422–31 (1947).
(39) Gucker, F. T., Jr., and O'Konski, C. T.: Chem. Revs. **44**, 373 (1949).
(40) Gurney, S. W., Williams, C. R., and Meigs, R. R.: J. Ind. Hyg. Toxicol. **20**, 24–35 (1938).
(41) Guyton, A. C.: J. Ind. Hyg. Toxicol. **28**, 133–41 (1946).
(42) Hamburger, M., Jr., Puck, T. T., and Robertson, O. H.: J. Infectious Diseases **76**, 208–15 (1945).
(43) Hatch, T. F., and Kindsvatter, V. H.: J. Ind. Hyg. Toxicol. **29**, 342–6 (1947).
(44) Heymann, B.: Zentr. ges. Hyg. Bakt. u. Immunitäts. **24**, 1 (1932).
(45) Johnson, I., and La Mer, V. K.: J. Am. Chem. Soc. **69**, 1184–92 (1947).
(46) Johnstone, H. F., Winsche, W. E., and Smith, L. W.: Chem. Revs. **44**, 353 (1949).
(47) La Mer, V. K., and Sinclair, D.: OSRD Report No. 1857 (September 29, 1943) and Part V of OSRD Rept. No. 865 (September 4, 1942).
(48) La Mer, V. K., and Barnes, M. D.: J. Colloid. Sci. **1**, 71–7 (1946).
(49) La Mer, V. K., and Hochberg, S.: Chem. Revs. **44**, 341 (1949).

(50) LANDAHL, H. D., AND BLACK, S.: J. Ind. Hyg. Toxicol. **29**, 269–77 (1947).
(51) LANDAHL, H. D.: Personal communication.
(52) LEHMANN, G.: J. Ind. Hyg. **17**, 37–40 (1935).
(53) LEMON, H. M.: Proc. Soc. Exptl. Biol. Med. **54**, 298–301 (1943).
(54) LISTER, J.: Brit. Med. J. **2**, 53–6, 101–2, 461–3, 515–17 (1868).
(55) LOOSLI, C. G., ROBERTSON, O. H., AND PUCK, T. T.: J. Infectious Diseases **72**, 142–53 (1943).
(56) LUCKIESH, M., HOLLADAY, L. L., AND TAYLOR, A. H.: Gen. Elec. Rev. **49**, 8–17 (1946).
(57) LUCKIESH, M.: *Applications of Germicidal, Erythemal and Infrared Energy.* D. Van Nostrand Company, Inc., New York (1946).
(58) MANSON, P.: J. Linnean Soc. London **14**, 304–11 (1879).
(59) MASPERO, SIR G.: *Life in Ancient Egypt and Assyria,* p. 119. Chapman and Hall, London (1891).
(60) MASTERMAN, A. T.: J. Hyg. **41**, 44–64 (1941).
(61) MAY, K. R.: J. Sci. Instruments **22**, 187–95 (1945).
(62) MIE, GUSTAV: Ann. Physik [4] **25**, 377–445 (1908).
(63) MIQUEL, P.: *Les organismes vivants de l'atmosphere.* Gauthier-Villars, Paris (1883).
(64) MOULTON, S., PUCK, T. T., AND LEMON, H. M.: Science **97**, 51–2 (1943).
(65) MUELLER, HANS: Personal communication.
(66) OSLER, SIR WILLIAM: *The Evolution of Modern Medicine.* Yale University Press, New Haven, Connecticut (1922).
(67) OWENS, J. S.: Proc. Roy. Soc. (London) **A101**, 18–37 (1922); J. Ind. Hyg. **4**, 522–34 (1922).
(68) PASTEUR, L.: *Memoire sur les corpuscules organismes qui existent dans l'atmosphere* (1862), *Oeuvres de Pasteur,* P. Vallery-Radot edition, Vol. 2, pp. 210–94. Masson et Cie, Paris (1922).
(69) PERRIN, F.: J. Chem. Phys. **10**, 415–27 (1942).
(70) PUCK, T. T.: J. Exptl. Med. **85**, 729–39 (1947).
(71) PUCK, T. T., ROBERTSON, O. H., AND LEMON, H. M.: J. Exptl. Med. **78**, 387–406 (1943).
(72) PULVERTAFT, R. J. V.: J. Hyg. **43**, 352–6 (1943–44).
(73) RIEDEL, G.: Kolloid-Z. **103**, 228–32 (1943).
(74) ROBERTSON, O. H., BIGG, E., PUCK, T. T., AND MILLER, B. F.: J. Exptl. Med. **75**, 593–610 (1942).
(75) ROBERTSON, O. H., LOOSLI, C. G., PUCK, T. T., BIGG, E., AND MILLER, B. F.: Science **94**, 612–13 (1941).
(76) ROOKS, R.: Am. J. Hyg. **30**, 7–10 (1939).
(77) ROSEBURY, T.: *Experimental Air-Borne Infection,* Microbiological Monographs. The Williams & Wilkins Company, Baltimore (1947).
(77a) ROSENBLATT, P., AND LA MER, V. K.: Phys. Rev. **70**, 385–95 (1946).
(78) SELL, W.: Forsch. Gebiete Ingenieurw., Forschungsheft **No. 347**, Berlin (1931).
(79) SONKIN, L. S.: J. Ind. Hyg. Toxicol. **28**, 269–72 (1946).
(80) STILLMAN, E. G., AND BRANCH, A.: J. Exptl. Med. **40**, 733–42 (1924).
(81) TOLMAN, R. C., GERKE, R. H., BROOKS, A. P., HERMAN, A. G., MULLIKEN, R. S., AND SMYTH, H. DE W.: J. Am. Chem. Soc. **41**, 575–87 (1919).
(82) TOLMAN, R. C., REYERSON, L. H., BROOKS, A. P., AND SMYTH, H. D.: J. Am. Chem. Soc. **41**, 587–9 (1919).
(83) TOLMAN, R. C., AND VLIET, E. B.: J. Am. Chem. Soc. **41**, 297–300 (1919).
(84) TOPLEY, W. W. C., AND WILSON, G. S.: *Principles of Bacteriology and Immunity,* revised by G. S. Wilson and A. A. Miles. The Williams & Wilkins Company, Baltimore (1936).
(85) TOURANGEAU, F. J., AND DRINKER, P.: J. Ind. Hyg. Toxicol. **19**, 53–7 (1937).
(86) TRILLAT, A.: Bull. acad. méd. (Paris) [3] **119**, 64–74 (1938).
(87) TWORT, C. C., BAKER, A. H., FINN, S. R., AND POWELL, E. O.: J. Hyg. **40**, 253–344 (1940).

(88) TWORT, C. C., AND BAKER, A. H.: J. Hyg. **44**, 451–63 (1945–46).
(89) TYNDALL, JOHN: *The Floating Matter of Air in Relation to Putrefaction and Infection.* D. Appleton-Century Company, New York (1882).
(90) TYNDALL, JOHN: Phil. Mag. [4] **37**, 384 (1869).
(91) WATSON, H. H.: Trans. Inst. Mining Metals (London) **46**, 155 (1936).
(92) WELLS, W. F.: Am. J. Pub. Health **23**, 58–9 (1933).
(93) WELLS, W. F.: Am. J. Hyg. **20**, 611–18 (1934).
(94) WELLS, W. F., AND STONE, W. R.: Am. J. Hyg. **20**, 619–27 (1934).
(95) VAN WIJK, A. M., AND PATTERSON, H. S.: J. Ind. Hyg. Toxicol. **22**, 31–5 (1940).
(96) WHYTLAW-GRAY, R., AND LOMAX, R.: *Smoke*, p. 186. Edward Arnold and Company, London (1932).
(97) WILSON, I. B., AND LA MER, V. K.: J. Ind. Hyg. Toxicol. **30**, 265–81 (1948).
(98) ZIMM, B. H., STEIN, R. S., AND DOTY, P.: Polymer Bull. **1**, 90–119 (1945).

APPENDIX A

B. globigii spores

For impinger studies, the spores were autoclaved 45 min. at 15 p.s.i. and sterility tests made. Suspensions were diluted with distilled water to give a final concentration of about 10^8 per cubic centimeter. For viability studies, autoclaving was omitted.

S. marcescens

This organism was grown for 18 hr. at room temperature in tryptic digest broth with 0.3 per cent dextrose added. Cultures were centrifuged, re-suspended in sterile distilled water, and allowed to stand overnight in the cold room. A second centrifugation was followed by re-suspension, 2–3 hr. standing, and final re-suspension in distilled water; if dead suspensions were desired, final re-suspension was in 0.5 per cent formaldehyde.

Total concentrations were determined by the method of direct count (84) in a Petroff-Hausser counting chamber on the day of use. Great care was taken to count individual organisms. Final concentrations were generally of the order of 10^8 per cubic centimeter.

The concentration of viable organisms was established by making appropriate dilutions, and counting colonies on pour plates.

APPENDIX B

Although direct determination of this ratio did not appear possible in terms of easily and rapidly measurable physical differences between living and dead organisms, indirect means of estimation did exist. Dunklin and Puck (25) have shown that at room temperature and humidity exceeding 60 per cent, the death rate of organisms in an aerosol is in terms of minutes and hours, not seconds. It therefore seemed reasonable to assume that under those atmospheric conditions viability would not change measurably as an organism passed from the atomizer to the collecting device. The time of transit is surely less than 3 sec. and probably less than 1 sec.

It remained to be shown that the viability of organisms was not reduced by atomization, although direct evidence is not now available. Supporting evidence can be obtained by determining viability ratios on the liquid suspension used for atomization both at the beginning of an experiment and on the residual fluid remaining in a reflux atomizer at its conclusion. In this type of atomizer larger droplets, many of them containing bacteria, return to the fluid reservoir. Consequently, some bacteria must be atomized more than once. If the process is harmful, viability of the bacterial suspension would decrease demonstrably with time. The absence of change would suggest that (*1*) atomization did not alter viability, and (*2*) indirectly that it was permissible to assume the same viability coefficient for both suspension and aerosol. The following measurements were necessary:

(*a*) The concentration of viable organisms was determined by the pour-plate method, using an incubation period of 48 hr. The temperature was 37°C. for *B. globigii* and about 25°C. for *S. marcescens*.

(*b*) Direct bacterial counts were made in a counting chamber. In addition to the usual count, the proportion of organisms occurring singly or in clumps was carefully observed. Moreover, the number of clumps containing two, three, or more organisms was separately recorded. Estimates of total bacterial concentration were computed in the customary manner.

Apparent viability coefficients presumably represent true coefficients only when a bacterial suspension is composed of single organisms. When clumps occur, the number of colonies appearing in plate cultures depends on the combined number of single particles and clumps. In the case of suspensions containing only single organisms and doublets, as with *S. marcescens*, it is possible to compute the viability ratios provided the proportion of organisms occurring singly and the apparent ratio are known. From these data, together with customary probability considerations, and the assumption that the probability of viability is the same for each cell whether it be single or in a doublet, one can readily derive the equation

$$V^2 - \frac{2V}{D} + \frac{2R}{D} = 0$$

in which B denotes the true and R the apparent viability ratio, and D the proportion of cells observed as doublets.

With *B. globigii* spores such treatment is not convenient, since more than four organisms frequently appear in a clump, and equations of higher degree result. Moreover, these spores are highly resistant and the assumption of 100 per cent viability is consistent with the observed number of colonies and the apparent viability ratios.

THE RELATIONSHIP BETWEEN PERFORMANCE AND CONSTITUTION OF PURE ORGANIC EXPLOSIVE COMPOUNDS

WARREN C. LOTHROP AND G. RICHARD HANDRICK[1]

Arthur D. Little, Inc., Cambridge, Massachusetts

Received May 5, 1948

I. INTRODUCTION

An intriguing problem of structural organic chemistry which has not been reviewed in the published literature is the relationship between the performance of pure explosive compounds and their constitution. Early attempts at developing a theory for predicting the power of explosive compounds have been almost entirely based on thermodynamic concepts. Berthelot (1) wrote one of the earliest and best expositions on this subject, but the relationship between chemical structure and performance was not considered, since publication was prior to most of the work in structural organic chemistry. Brunswig (4), Moser (11), Marshall (10), Stettbacher (22), and others (6, 13) have examined this field of chemistry in much the same manner as Berthelot.

Some investigators (3, 10, 18, 22) have implied the value of oxygen balance, a property which can be calculated from empirical formulas, but none has stressed its fundamental applicability. Although the idea has been implicit in the thinking of those familiar with the field, there are few documented publications of it aside from a study of dynamites (14) and one by Roth (20) on mixtures of tetranitromethane and nitrobenzene or nitroglycerin.

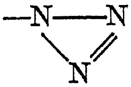
A single effort to develop a relationship between performance and structural composition has been made. This is that of Pletz (15), who proposed a theory of "explosophores" and "auxoexploses," analogous to the Witt theory of chromophores and auxochromes and to the Ehrlich theory of toxophores and autotoxes. According to this theory the explosive properties of any given substance depend upon the presence of definite structural groupings, called explosophores. the auxoexploses fortify or modify the explosive properties brought about by the explosophore.

On the basis of this theory Pletz examined eleven classes of organic compounds and subdivided all explosives into eight classes containing as "explosophores":

I. —NO_2 and —ONO_2 groups connected to organic or inorganic radicals (for example, HNO_3, $C(NO_2)_4$, and $C_6H_3(NO_2)_3$).

---

[1] The authors are greatly indebted to Colonel C. H. M. Roberts and Dr. L. R. Littleton of the Office of the Chief of Ordnance for assistance and encouragement in connection with the initiation and progress of the investigation, and to many others expert in the field who have given freely of their time and knowledge. The work was done under Contract W-19-020-ORD-6436 with the Office of the Chief of Ordnance and has been released by the War Department for publication.

II. $-N{=}N-$ and  groups, connected to inorganic or organic radicals (for example, PbN_6 and CH_3N_3).
III. $-NX_2$ group (for example, NCl_3 and $RNCl_2$).
IV. $C{=}N-$ group (for example, HONC and $Hg(ONC)_2$).
V. $-OClO_2$ and $-OClO_3$ groups, connected to inorganic or organic radicals (for example, $KClO_3$, $KClO_4$, and certain organic perchlorates).
VI. $-O-O-$ and $-O-O-O-$ groups connected to inorganic or organic radicals (for example, peroxides and ozonides).
VII. $-C{\equiv}C-$ group (for example, acetylene or its metallic derivatives).
VIII. A metal atom connected by an unstable bond to the carbon of certain organic radicals (for example, organic compounds of mercury, thallium, and lead).

While the arrangement of Pletz embraces the whole explosive field in a purely empirical fashion, the distinction between the terms "explosophore" and "auxoexplose" is vague, poorly developed, and of little practical usefulness.

He and other authors were not fortunate in having before them full information on a great variety of explosive compounds, for, although a considerable body of data had been amassed by workers throughout the world prior to the Second World War, a large fraction of it was buried in confidential government or industry files. That which was published appeared in relatively obscure journals and was not infrequently of questionable authenticity. Before any study such as the one to be described here could be undertaken, it was necessary that this information be assembled in usable form. This tremendous task was the subject of an extremely successful wartime project conducted by Dr. A. H. Blatt. It can be said without qualification that his compilation alone has made possible the presently reported treatment.

The object of the work has been the collation and classification of information through relationships which exist between organic structure and the explosive properties of greatest interest in military applications. These properties have been judged to be power (the work capability or strength of an explosive) and brisance (the shattering effect connected with the speed of detonation). Other properties are also mentioned briefly. The compounds considered are chiefly "high explosives," as contrasted with initiating explosives and commercial compositions.

Physical chemistry, by the use of standard thermodynamic concepts, has successfully treated the subject quantitatively through the additivity of bond energies and the use of standard calorimetry. This aspect is touched on lightly here, for it is hoped that it may someday be published fully by those who have been so successful with it.

It is believed that using the concepts explained one should be better guided in the synthesis of new and powerful high explosives and the prediction of their explosive properties. Whether any compound would thereby be discovered greatly superior to the excellent explosives now in use is problematical. For military purposes, moreover, it is likely that items other than explosive performance would be governing, such as economics, effectiveness in special uses, and difficulty of synthesis.

II. CONSTITUTION AND EXPLOSIVE POWER

An explosive compound can be looked upon as a chemical combination of atoms in a metastable state which spontaneously, on passage of a detonating wave, rearrange themselves to give simple products with a concurrent liberation of large volumes of gas and considerable energy, mainly in the form of heat. For purposes of this treatment the detonation products may be considered as consisting of the free elements or their several oxides. They are chiefly water, carbon or its oxides, and nitrogen, although traces of other substances are also known to be formed, such as hydrogen cyanide and methane. Other elements are occasionally encountered, such as chlorine, sulfur, lead, or mercury, but organic high explosives in the main are composed of combinations of only four elements: carbon, hydrogen, oxygen, and nitrogen.

Since combustion is the process taking place, a simple computable term should be of value in stating the degree to which oxidation is possible. For this purpose the expression known as *oxygen balance* is useful. This is given as the percentage of oxygen required for complete conversion of the carbon and hydrogen present to carbon dioxide and water. For any compound containing x atoms of carbon, y atoms of hydrogen, and z atoms of oxygen, the oxygen balance = $\left[-1600\left(2x + \frac{y}{2} - z\right)\right]\Big/$molecular weight. Thus an explosive having perfect balance to yield carbon dioxide and water has zero balance, one lacking sufficient oxygen has a negative balance, and one containing excess oxygen has a positive balance. It is noted that with very few exceptions explosives are of decidedly negative balance. This method of computation makes no distinction between oxygen already bound to carbon or hydrogen and that bound to nitrogen. The latter might be termed "effective" oxygen and differs from the former in that it is loosely held and available for subsequent combustion.

A term which will describe power quantitatively is more difficult to devise, although it has been successfully accomplished by unpublished thermodynamic treatment. Two of the most commonly employed experimental tests designed for power measurement are the so-called Trauzl lead block expansion test (12, 16, 25) and the ballistic mortar test (24). The former is a classical method and depends on determining the expansion caused in a block of "pure" lead when a standard weight of explosive is detonated in a specified cavity under controlled confinement. It is customarily expressed in this country in terms of TNT, whose arbitrary value is given as 100. Whether the lead block test is a measure of power alone or of an uncertain combination of power and brisance is a subject of contention but not of great importance, since it can be shown that power and brisance appear to vary concomitantly. Caution should govern the use of the lead block values, however, as an absolute measure of power.

The second method of testing for power is by the ballistic mortar, in which a standard explosive charge is fired in a heavy pendulum mortar and the angle of recoil is measured. This method can also be related to TNT as a standard of 100, and is considered to be uncomplicated by brisance. It is probably safer to use ballistic mortar values in direct comparisons of explosive power, for "the high explosive instead of being caused to do work against a varying and unknown resistance as in the lead block is fired in a suitable type of gun, so that the work done by expansion of the products of detonation is measured in absolute mechanical units" (26).

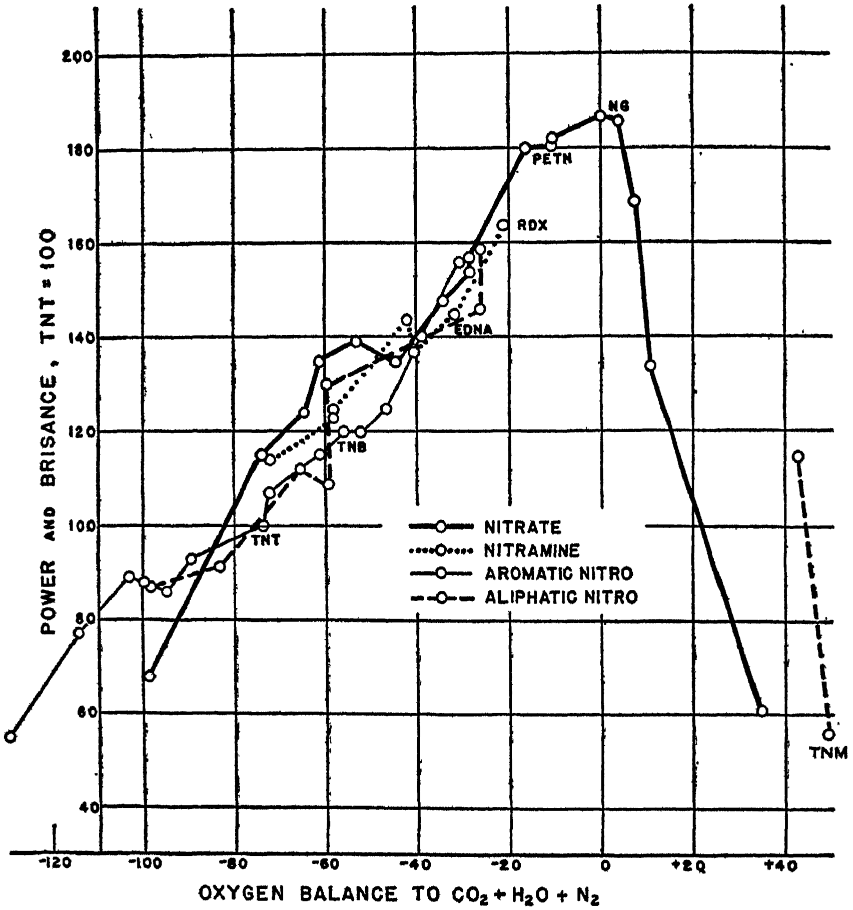


FIG. 1. Effect of oxygen balance on power and brisance of explosives compared with TNT in the lead block expansion test. Primary plosophores.

The results of both methods appear to be equally useful and to yield the same conclusions. The lead block data, however, are much more extensive.

If one plots the lead block values for known organic explosives against their indicated oxygen balance, one is confronted with a mass of points scattered at random over the whole graph. This apparent confusion can be very quickly resolved by the use of the term *plosophore.*[2] By analogy with color chemistry, this has been defined as a group of atoms which on substitution into a hydrocarbon is capable of forming an explosive compound. Inspection of the substituent groups which can function in this way indicates at once that there are two classes of plosophores, differing sharply in effectiveness and consistency in producing power and hence properly called *primary* and *secondary.*

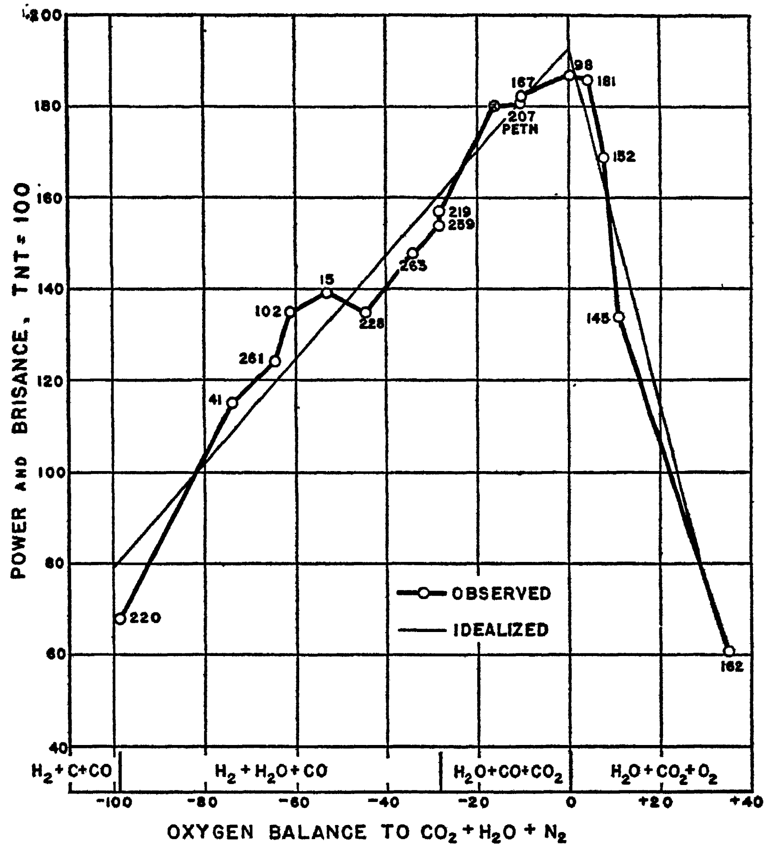


FIG. 2. Effect of oxygen balance on power and brisance of explosives compared with TNT in the lead block expansion test. Primary plosophores: nitrate.

Primary plosophores include the following groups: nitrate ester, aromatic nitro, aliphatic nitro, and nitramine. The secondary plosophores make up the remainder and include such groups as azo, azide, nitroso, peroxide, ozonide, perchlorate, etc.

In figure 1[3] all compounds containing carbon, hydrogen, and any one type of the four primary plosophores are compared in lead block value and oxygen balance. Considering that all the data available have been used without regard to quality of source and that some of them have been interpolated from ballistic mortar results by the use of calculated factors, the family of curves which results is remarkably consistent. The conclusion is inescapable that power is directly related to oxygen balance for primary plosophoric compounds, that it is at a maximum in compounds whose balance is close to zero, and that the four types of explosives are roughly equal to each other at given balance although their indicated maxima may not exactly coincide (3, 14).

---

[2] The present terms were derived independently of Pletz (15). Since their definition and use differ from his terms, a different nomenclature is felt to be justified.

[3] The abbreviations on figure 1 for the more important high explosives stand for the scientific names as follows: TNT = 2,4,6-trinitrotoluene; TNB = 1,3,5-trinitrobenzene; TNM = tetranitromethane; RDX = cyclotrimethylenetrinitramine; PETN = pentaerythritol tetranitrate; NG = nitroglycerin; EDNA = ethylenedinitramine.

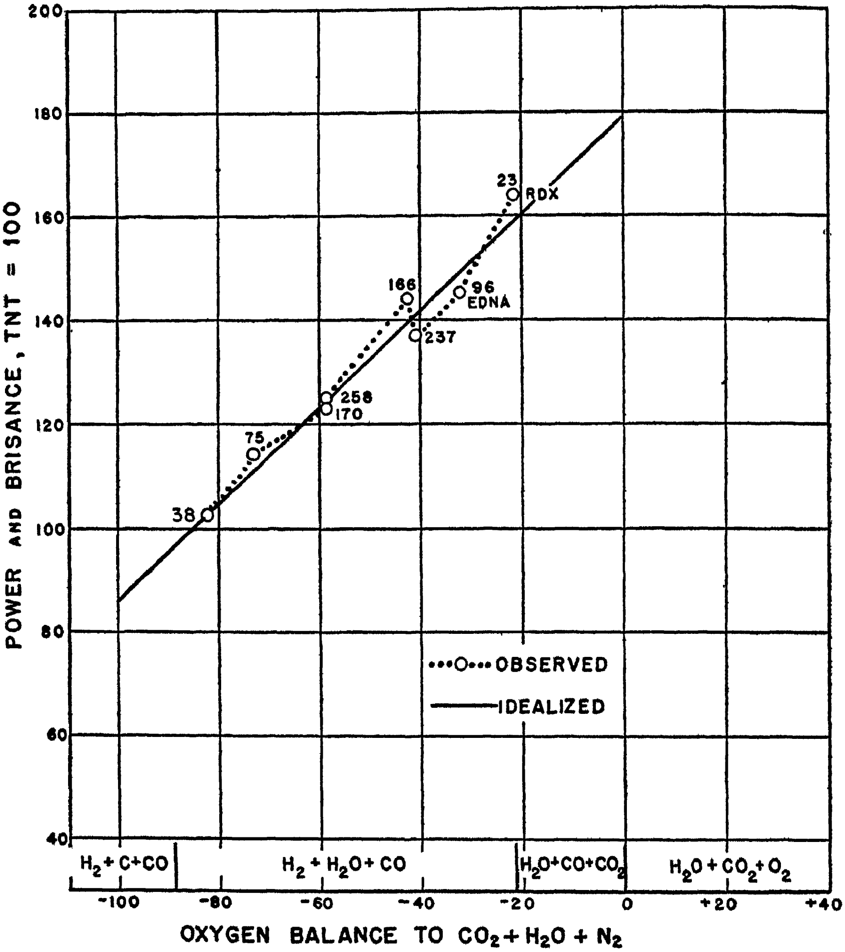


FIG. 3. Effect of oxygen balance on power and brisance of explosives compared with TNT in the lead block expansion test. Primary plosophores: nitramine.

Treatment of each plosophoric series separately and use of the least square method for the derivation of idealized straight lines to express the data yield the results plotted in figures 2 to 5. It is noted that these lines do not coincide, owing to a real difference or the inadequacy of the data. In these and other figures the data have been conveniently summarized by the drawing of idealized curves. The authors consider them as useful generalizations of no great accuracy. When straight lines are drawn they are shown to reach a maximum at zero oxygen balance, but in practice it appears likely that the true maximum will be found in the region of negative balance, although close to zero. Because of the scant data in this region, it is impossible to locate the point exactly. The figures also bear numbers which refer to the individual compounds listed in table 1. Some unidentified points are reported where release of information could not be secured. It is observed at once that only the nitrate ester series is reasonably complete, with the other three series lacking examples in the interesting region of oxygen balance lying between −20 per cent and ±0 per cent. At the foot of figures 2 to 5 are approximate indications of the assumed products of detonation as they vary with oxygen balance.

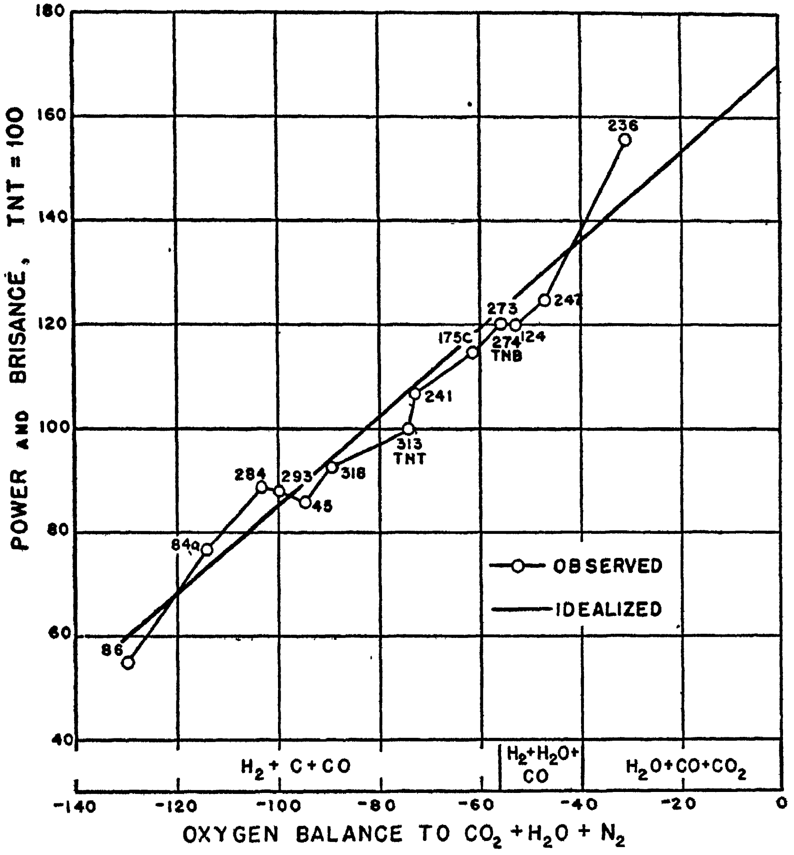


FIG. 4. Effect of oxygen balance on power and brisance of explosives compared with TNT in the lead block expansion test. Primary plosophores: aromatic nitro.

Having established the desirability of primary plosophores as power-producing groups, it is of interest to learn whether the presence in a single molecule of more than one type of such groups has an effect. Such molecules may be termed *hybrid*.

Of the six possible hybrids of two primary plosophores, four are represented by available data; in addition, examples of molecules are found containing three different primary plosophores. All these hybrids are given in figure 6, together

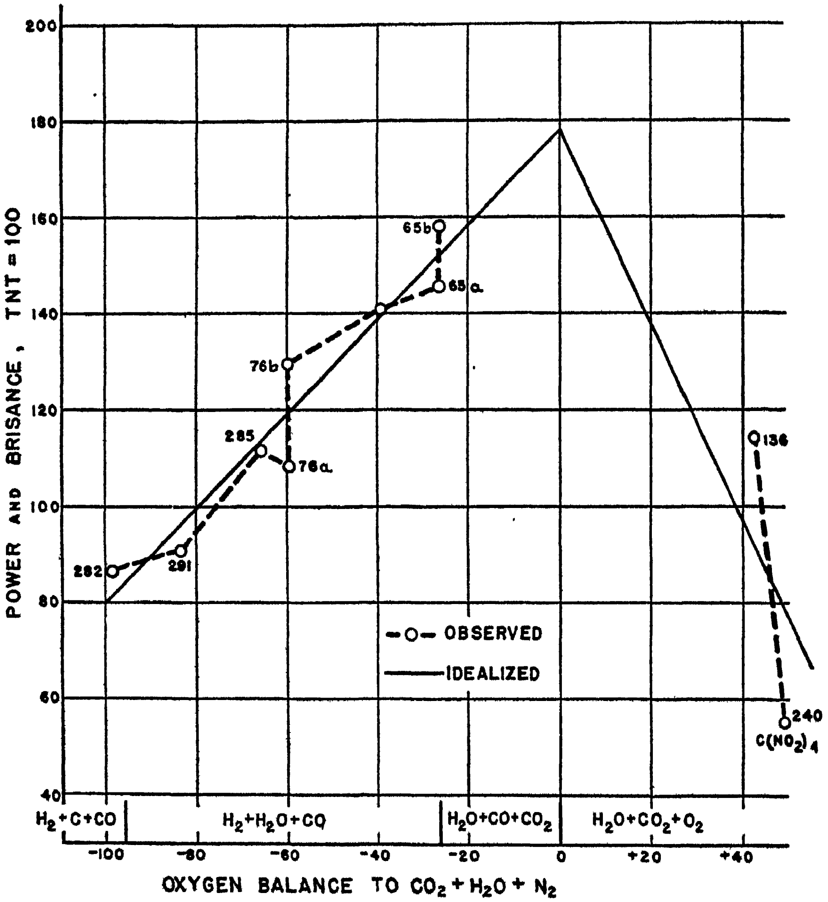


FIG. 5. Effect of oxygen balance on power and brisance of explosives compared with TNT in the lead block expansion test. Primary plosophores: aliphatic nitro.

with the single idealized line for the performance of the pure primary plosophores. It is interesting to note how abruptly the data dwindle in passing beyond an oxygen balance of −20 per cent, thus duplicating the situation already noted in figure 1. Nevertheless, the result is obvious: hybrids are the equivalent of pure types, although they tend to exhibit somewhat greater variability. In any synthetic program on high explosives this equivalence of hybrids will be of importance, because it extends enormously the synthetic possibilities.

The very clean-cut relationship between power and oxygen balance vanishes

TABLE 1

Index of compounds

COMPOUND NUMBER	
3	Ammonium nitrate
5	Ammonium picrate
9	Azidoethyl alcohol
10	Azidoethyl nitrate
15	1,3-Butylene glycol dinitrate
17	β-Chloroethyl nitrate
18	Cyanoethyl nitrate
20	Cyanuric triazide
23	Cyclotrimethylenetrinitramine
24	Cyclotrimethylenetrinitrosamine
28	Diazodinitrophenol
32a	1,2-Dichloro-3,5-dinitrobenzene
32b	1,3-Dichloro-4,6-dinitrobenzene
33	Dicyandiamidine perchlorate
34	Di(β,γ-dihydroxypropyl)oxamide tetranitrate
37	Diethylene glycol dinitrate
38	*N*,*N′*-Dimethyl-*N*,*N′*-dinitroethylenediamine
41	2,2-Dimethylpropane-1,3-diol dinitrate
43	2,4-Dinitroaniline
44	2,4-Dinitroanisole
45	*m*-Dinitrobenzene
46	2,4-Dinitrobenzyl nitrate
50	Dinitrodiethyloxamide
51	1,4-Dinitro-3,6-diketopiperazine
52	2,3-Dinitro-2,3-dimethylbutane
53	1,4-Dinitro-2,5-dimethyl-3,6-diketopiperazine
57	1,4-Dinitro-2,5-dimethylolbenzene dinitrate
58	*N*,*N′*-Dinitrodimethyloxamide
59	1,3-Dinitro-2,2-dimethylpropane
60	Dinitrodimethylsulfamide
62	1,3-Dinitro-4,5-dinitrosobenzene
64	Dinitrodi(β-nitroxyethyl)sulfamide
65a	1,1-Dinitroethane
65b	1,2-Dinitroethane
66	*N*,*N′*-Dinitroethyleneurea
67	2,5-Dinitrofuran
71	2,4-Dinitrophenol
72	β-(2,4-Dinitro)phenoxyethyl nitrate
73	3,5-Dinitrophenylnitromethane
74	2-(3′,5′-Dinitrophenyl)-2-nitro-1,3-propanediol dinitrate
75	1,4-Dinitropiperazine
76a	1,1-Dinitropropane
76b	2,2-Dinitropropane
80	2,4-Dinitrosoresorcinol
81	1,4-Dinitro-1,1,4,4-tetramethylolbutane tetranitrate

COMPOUND NUMBER	
82	1,5-Dinitro-1,1,5,5-tetramethylolpentane tetranitrate
84a	2,4-Dinitrotoluene
84b	2,6-Dinitrotoluene
86	4,6-Dinitro-*m*-xylene
89	Dipentaerythritol hexanitrate
93	Erythritol tetranitrate
94	Ethanolamine dinitrate
95	Ethylenediamine dinitrate
96	Ethylenedinitramine
98	Ethylene glycol dinitrate
102	Ethyl nitrate
105	Glycerol 1,2-dinitrate
107	Glycerol α-2,4-dinitrophenyl ether dinitrate
108	Glycerol α,α′-hexanitrodiphenyl ether β-nitrate
109	Glycerol monochlorohydrin dinitrate
111	Glycerol monolactate trinitrate
112	Glycerol α-monomethyl ether dinitrate
113	Glycerol α-2,4,6-trinitrophenyl ether dinitrate
114	Guanidine nitrate
115	Guanidine perchlorate
116	Guanylnitrourea
118	Hexamethylenetriperoxidediamine
120	Hexamine dinitrate
123	Hexanitroazobenzene
124	Hexanitrobiphenyl
125	Hexanitrodiphenylamine
127	β-(Hexanitrodiphenylamino)ethyl nitrate
128	2,4,6,2′,4′,6′-Hexanitrodiphenyl ether
130	*N*,*N*′-(Hexanitrodiphenyl)ethylenedinitramine
132	Hexanitrodiphenylmethylamine
133	Hexanitrodiphenyl sulfide
134	Hexanitrodiphenyl sulfone
135	Hexanitrodiphenylurea
136	Hexanitroethane
137	Hexanitrohydrazobenzene
138	Hexanitroöxanilide
143	Hydrazine perchlorate
145	Inositol hexanitrate
148	Lactose octanitrate
150	Lead styphnate
151	Maltose octanitrate
152	Mannitol hexanitrate
154	Mercuric fulminate
155	Methazonic acid

COMPOUND NUMBER	NAME
156	Methylamine nitrate
162	Methyleneglycol dinitrate
163	*N*-Methylethylenedinitramine
166	Methylnitramine
167	Methyl nitrate
168	2-Methyl-2-nitro-1,3-propanediol dinitrate
170	Trimethylene-1,3-dinitramine
172	Pentaerythritol diglycolate tetranitrate
175a	2,4-Dinitrophenylazonitromethane
175b	2,4,6,2′,4′,6′-Hexanitro-3,5,3′,5′-tetrahydroxybiphenyl
175c	Tetranitro-*m*-xylene
178	β-Nitroethanol
181	Nitroglycerin
182	Nitroguanidine
185	Nitromethane
186	1-Nitro-5-methylhydantoin
192	*m*-Nitrophenyldinitromethane
193	Nitrophenylglycol dinitrate
194	1-(*o*-Nitrophenyl)-2-nitroethyl nitrate
197	Nitrosoguanidine
200	Nitrourea
205	Pentaerythritol dimethyl ether dinitrate
206	Pentaerythritol monomethyl ether trinitrate
207	Pentaerythritol tetranitrate
210	Pentanitrodiphenyl ether
212a	Ethyl perchlorate
212b	Ammonium perchlorate
212c	Ethylenediamine diperchlorate
212d	1,2-Propylenediamine diperchlorate
212e	1,3-Propylenediamine diperchlorate
216	Picric acid
217	Picryl chloride
218	Picryl fluoride
219	1,2-Propanediol dinitrate
220	Propyl nitrate
223	Sucrose octanitrate
225	Tetracene (1-guanyl-4-nitrosaminoguanyltetrazene)
228	Tetramethylolcyclohexanol pentanitrate
229	Tetramethylolcyclohexanone tetranitrate
231	Tetramethylolcyclopentanone tetranitrate
233	2,3,4,6-Tetranitroaniline
234	2,3,5,6-Tetranitroanisole
236	1,2,3,5-Tetranitrobenzene
237	Tetranitrodiglycerin
240	Tetranitromethane

TABLE 1—*Concluded*

COMPOUND NUMBER	
241	1,3,6,8-Tetranitronaphthalene
243	2,3,4,6-Tetranitrophenylmethylnitramine
247	2,3,4,6-Tetranitrotoluene
248	Tetra(β-nitroxyethyl)ammonium nitrate
253	Tetryl
258	Trimethylene-1,2-dinitramine
259	Trimethylene glycol dinitrate
261	1,1,1-Trimethylolbutane trinitrate
262	1,1,1-Trimethylol-2-chloroethane trinitrate
263	1,1,1-Trimethylolethane trinitrate
264	Trimethylolnitromethane trinitrate
270	2,4,6-Trinitroaniline
271	2,4,6-Trinitroanisole
272	2,4,6-Trinitrobenzaldehyde
273	1,2,4-Trinitrobenzene
274	1,3,5-Trinitrobenzene
275	2,4,6-Trinitrobenzoic acid
276	2,4,6-Trinitrobenzyl nitrate
277	2,4,6-Trinitro-*m*-cresol
279	2,4,6-Trinitro-1,3-di(methylnitramino)benzene
282	2,2,3-Trinitro-3-ethylpentane
283	Trinitromelamine
284	Trinitromesitylene
285	2,3,3-Trinitro-2-methylbutane
290	2,4,6-Trinitro-3-methylnitraminotoluene
291	2,2,3-Trinitro-3-methylpentane
292	2,4,6-Trinitromonomethylaniline
293	1,3,8-Trinitronaphthalene
294	2,4,6-Trinitrophenetole
295	2,4,6-Trinitrophenoxyethyl nitrate
296	2,4,6-Trinitrophenyl azide
297	2,4,6-Trinitrophenylbutylnitramine
300	2,4,6-Trinitrophenylethylnitramine
305	2,4,6-Trinitrophenyl(1-methylol)propylnitramine nitrate
307	β-(2,4,6-Trinitrophenylnitramino)ethyl nitrate
310	2,4,6-Trinitroresorcinol
313	2,4,6-Trinitrotoluene
314	Trinitrotoluenes
315	2,4,6-Trinitro-1,3,5-triazidobenzene
316	2,4,6-Trinitro-1,3,5-tris(methylnitramino)benzene
318	1,3,6-Trinitro-*m*-xylene
320	Urea nitrate

when one considers secondary plosophores. In figure 7 this is done by comparing the data for these compounds with a single idealized curve for the four primary types obtained by averaging the four separate curves of figures 2 to 5. It is no longer possible to say that increased oxygen balance is accompanied by increased power. This may on occasion be the case, but there is no assurance of it. Further, with few exceptions, secondary plosophores do not significantly exceed primary plosophores as power-producing substituents. Therefore, it can be concluded that secondary plosophores do not recommend themselves as a class in the synthesis of powerful organic explosives. This is not to say that they are useless, for they often impart desirable qualities of another kind as in forming initiating and detonating compounds, but if power and brisance alone are the criteria, the secondary plosophores are to be avoided.

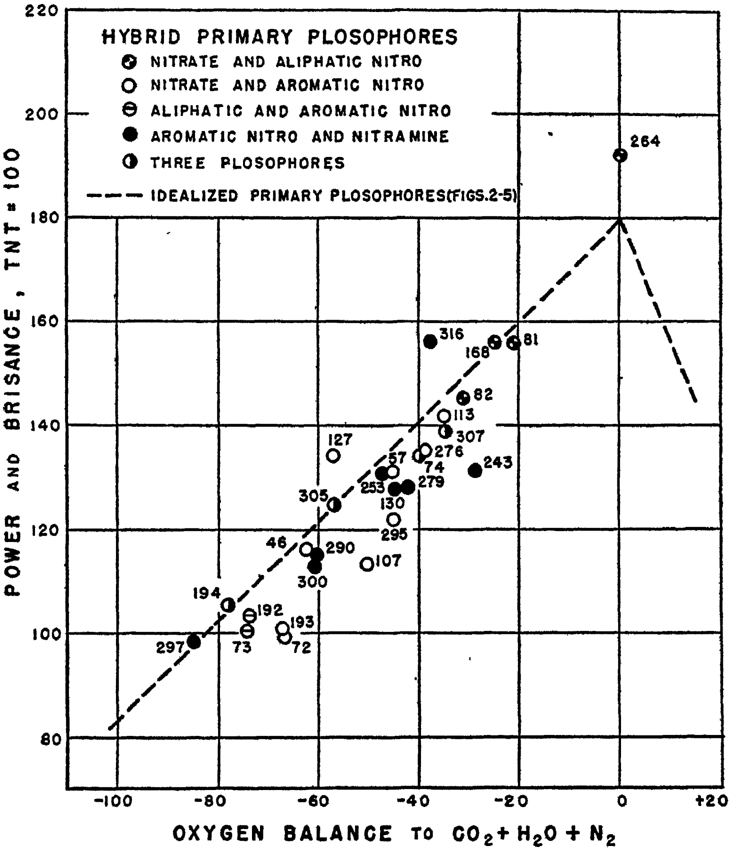


FIG. 6. Effect of oxygen balance on power and brisance of explosives compared with TNT in the lead block expansion test. Hybrid primary plosophoric compounds.

A very common feature of explosives is the presence in them of a large variety of substituent groups which are not plosophoric, since they do not in themselves produce explosive molecules, but which may be expected to alter the explosive properties in the same way that an auxochromic group is found to vary the intensity or shade of a dye. Such substituents may be designated *auxoplosive.*[4] Etymologists may object, because the name implies that such groups will always *aid* the formation of powerful explosives rather than *alter* their performance. Nevertheless, the term is retained in the interests of analogy and it is not believed that any serious confusion will arise. As examples of such groups, one may take hydroxyl, carbonyl, chloride, sulfide, ether, amino, etc. Not infrequently the presence of these groupings affects oxygen balance in a favorable manner, and it is of interest to see whether a concurrent improvement in power is produced. The data in figures 8 and 9 speak for themselves. With hardly an exception auxoplosive groups are shown to be detrimental to power whether or not they appear to improve oxygen balance. This is in agreement with the observation made earlier about "effective" oxygen. It is concluded, therefore, that the chief

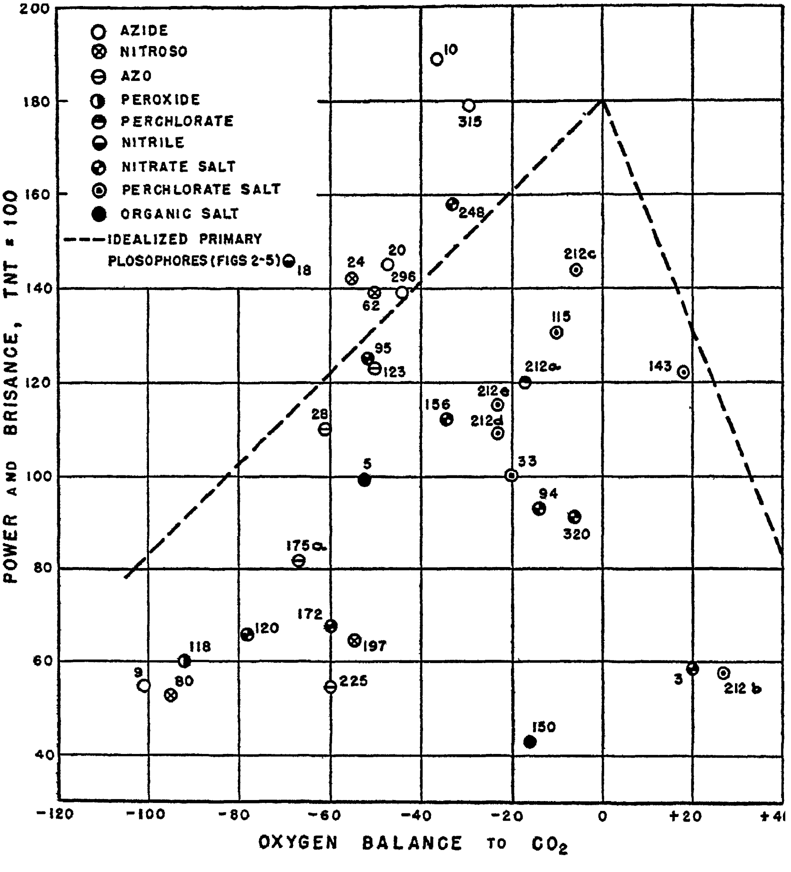


FIG. 7. Effect of oxygen balance on power and brisance of explosives compared with TNT in the lead block expansion test. Secondary plosophores.

---

[4] Pletz (15) uses the word "auxoexplose."

justification for the presence of such groups in powerful explosives is expediency in synthesis.

From the above discussion we can now say that power and brisance as measured in the Trauzl lead block are at a maximum for carbon and hydrogen compounds which bear as substituents only primary plosophores and are at or close to zero

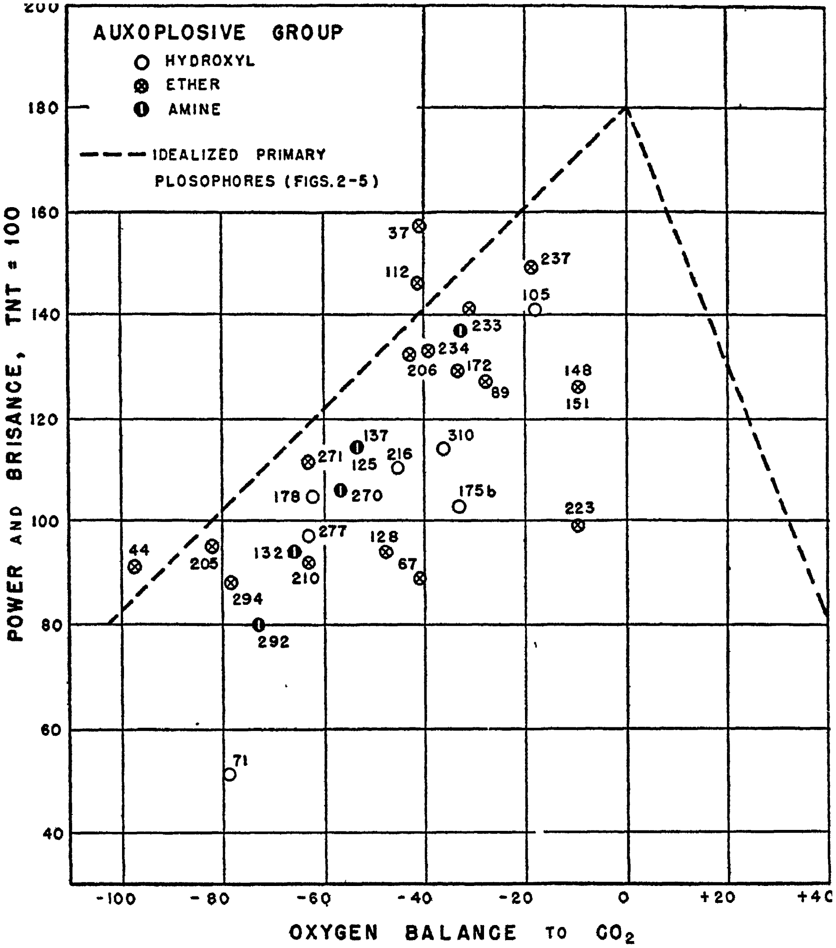


Fig. 8. Effect of auxoplosive groups on power and brisance of primary plosophoric compounds compared with TNT in the lead block expansion test.

oxygen balance. The presence of secondary plosophores and auxoplosive groups is not desirable, although on occasion these can be tolerated in the interests of increasing sensitivity (as in initiators) or of synthetic ease. Hybrid molecules may be classed as equivalent to pure plosophoric types.

Ballistic mortar data yield exactly similar conclusions, as is shown in figures 10 and 11. In these figures it is remarked that the observed power in terms of TNT is usually much lower than in the corresponding lead block graphs. Thus the apparent maximum at the zero balance point is about 160 instead of 180, —a value more comparable to actual experience in use. This points up the earlier recommendation that one should place less significance on an absolute lead block value and should use ballistic mortar results if a comparison of the relative powers of explosives is to be made.

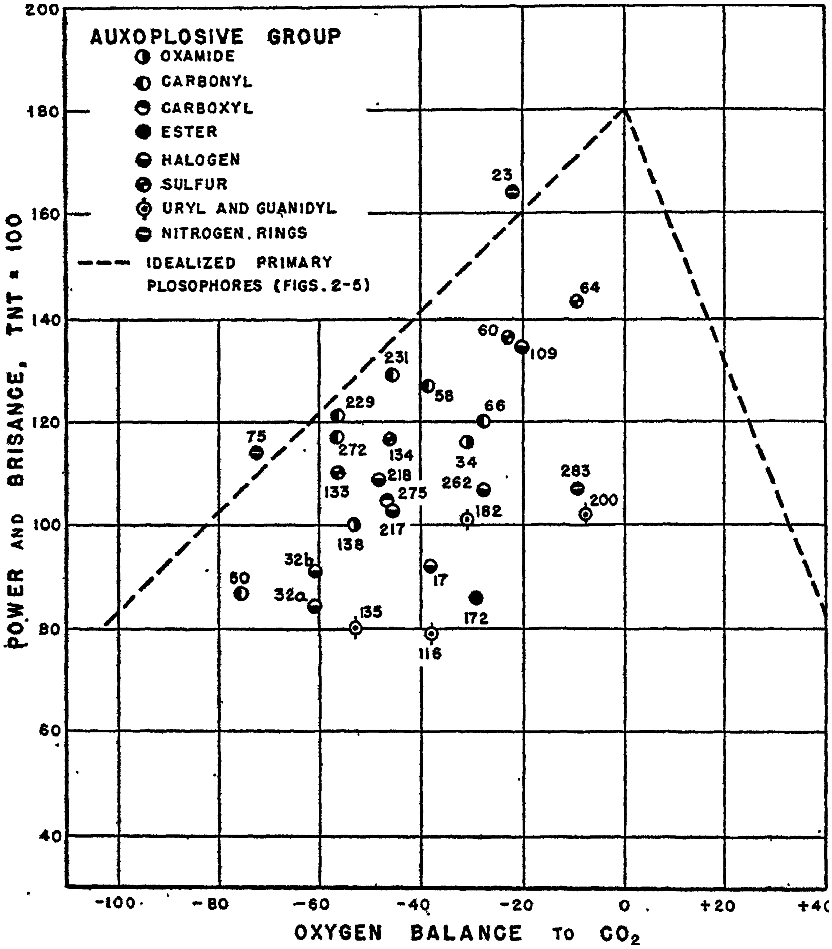


Fig. 9. Effect of auxoplosive groups on power and brisance of primary plosophoric compounds compared with TNT in the lead block expansion test.

Lastly, by either power scheme it can be shown that the effect of positional isomerism causes variations in power which are real but small (i.e., probably less than 10 per cent), while mere molecular size, as would be expected, is of no consequence.

III. CONSTITUTION AND BRISANCE

Employing the concepts of plosophoric and auxoplosive groups one can treat the separate problem of brisance in an analogous manner, although the paucity of data is a serious handicap. This is due chiefly to the lack of a truly satisfactory test which can measure the shattering effect uncomplicated by power. An early method which has lost some favor in recent years is the sand crush test (23). In this procedure a fixed weight of standard sand of known mesh size is enclosed in a bomb and subjected to the shattering action of a standard charge of explosive. The sand crushed is measured by resieving, and the quantities are compared to TNT, as in the case of a self-consistent set of data received by the authors from Picatinny Arsenal. When these data are plotted using the procedures so successful with power, a graph (figure 12) results in which sand crush values for primary plosophores ascend steadily with increasing oxygen balance and display an apparent maximum at the zero point. Secondary and auxoplosive compounds are seen to be inferior.

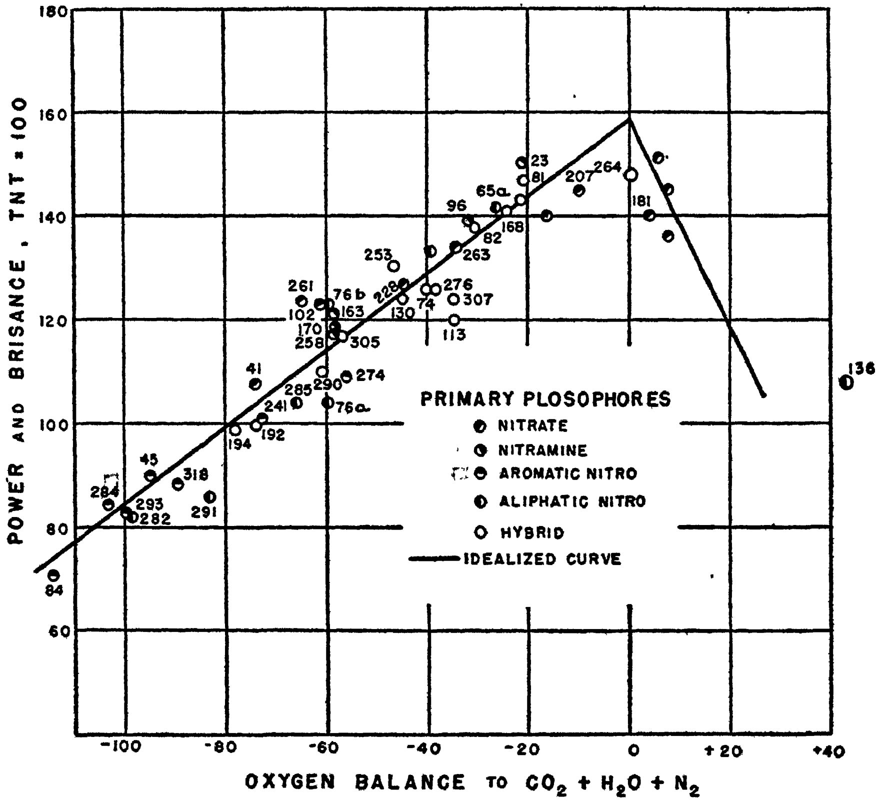


FIG. 10. Effect of oxygen balance on power and brisance of explosives compared with TNT in the ballistic mortar test. Primary and hybrid plosophores.

Less ambiguous but more difficult of quantitative analysis are tests measuring the degree of compression caused in standard lead and copper cylinders by the detonation of a standard charge of explosives placed on top, or tests measuring the depth of penetration of a stack of standard lead sheets when similarly attacked (5).

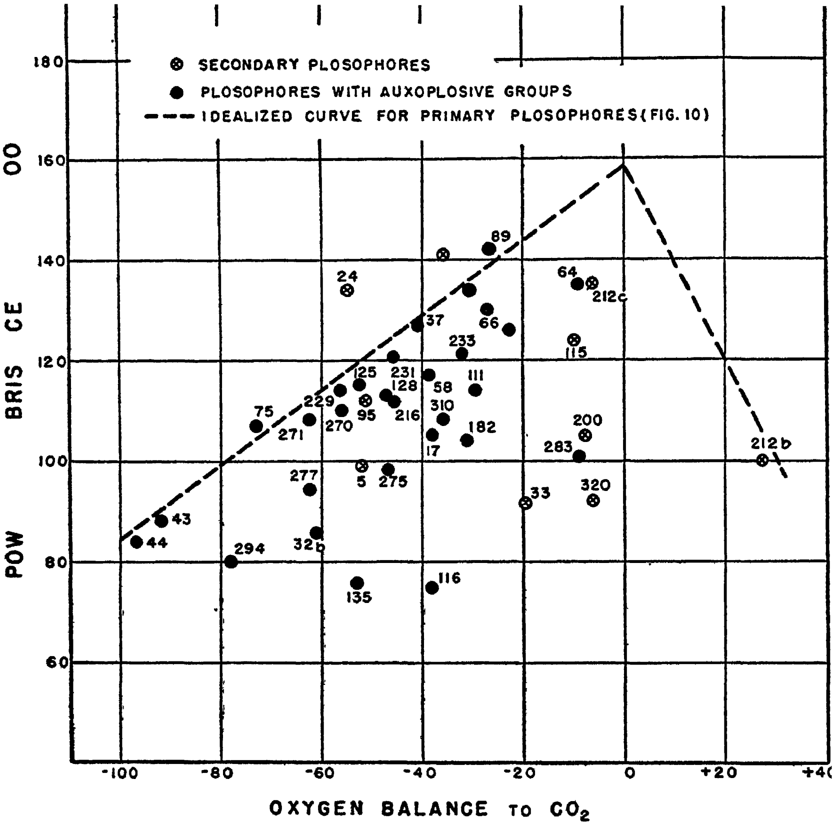
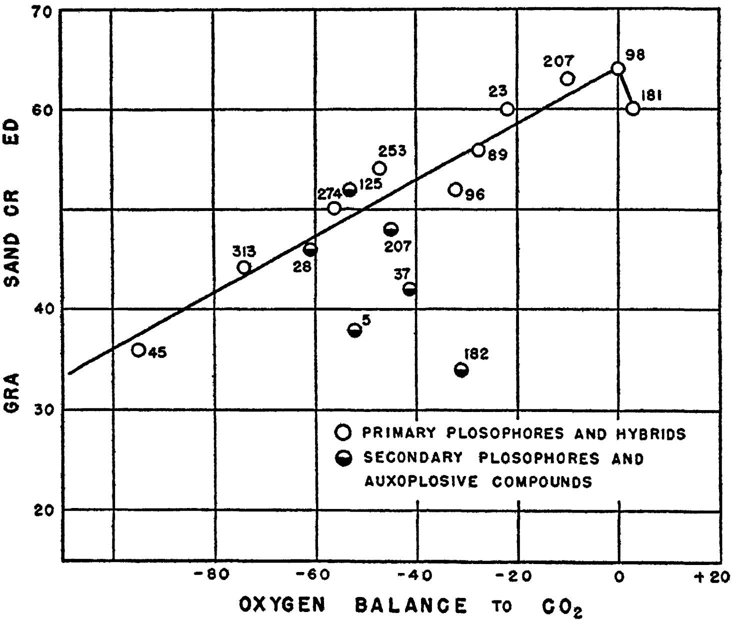


Fig. 11. Effect of oxygen balance on power and brisance of explosives compared with TNT in the ballistic mortar test. Secondary plosophoric and auxoplosive groups.



Interpretation of such tests is quite subjective, since the appearance of the damage caused is considered significant as well as the depth of penetration. An attempt to circumvent this feature has been made in which a steel plate supported annularly is substituted for lead. The result is a dent rather than a rupture, with the depth of dent being the measure of brisance. Graphical expression of the few data at hand has given results identical with sand crush, thus implying that brisance and power vary simultaneously and in the same direction. It is interesting to note in passing that all plate dent values when plotted against the corresponding ballistic mortar values give a straight line regardless of the presence of plosophoric or auxoplosive substituents. These observations confirm the judgment that power and brisance are twin qualities and that no false conclusions will result from the use of any single test so long as the data are self-consistent.

IV. CONSTITUTION AND OTHER EXPLOSIVE PROPERTIES

A treatment using the same concepts when applied to the properties of detonation velocity and sensitivity yields similar conclusions with somewhat less assurance. In figure 13 observed velocities for all types of explosives are plotted against oxygen balance at a constant loading density of 1.50. In cases where the velocity was measured at a density other than 1.50, a corrected value was obtained by use of the empirical formula $Dd_2 = Dd_1 + 3770(d_2 - d_1)$. There appears to be a drift toward higher velocities as the oxygen balance improves, but the exceptions are numerous. An idealized line has been drawn as a convenience.

It can also be demonstrated that the sensitivity of primary plosophoric and hybrid compounds increases rapidly with increasing oxygen balance, to reach a maximum at the zero balance point. There appears to be a less rapid falling off, however, in the region of positive balance than in the case of power.

The very interesting and valuable piece of work by Roth (20) fully supports all the relationships developed here and in the preceding discussion. Using homogeneous mixtures of tetranitromethane (a liquid of very high oxygen balance) and nitrobenzene (a liquid of very low oxygen balance), the whole range of oxygen balance was investigated and it was shown that power (lead block), detonation velocity, brisance (lead plate), and sensitivity all reach their maximum values at or very close to the point of zero oxygen balance. Of these qualities only sensitivity failed to diminish in the region of positive balance. The implications of these observations for military use are unfortunate, in that the most powerful explosives will also be expected to be the most sensitive.

On the same sensitivity scale, auxoplosive compounds are found to be slightly more variable than pure plosophoric types but otherwise similar. It is among the secondary plosophores that one encounters unpredictable behavior. Here the effect of different chemical groupings is a real one and is superimposed on that of oxygen balance, so that sensitivity is tied up to a variable and large degree with the specific identity of the substituent secondary plosophore. Azides are notably sensitive examples and hence find usefulness as initiators.

Although it would be particularly desirable to relate chemical stability to other parameters, it has not proved possible to do so except in a very rough manner, as has been done in the past. *A priori* it can be said that nitrate esters are generally unstable, that primary nitramines are less stable than secondary, that aliphatic nitro compounds are most stable when they lack enolizable hydrogen, etc. These observations have been made before and no available quantitative data appear sufficient to improve upon them or to reduce them to more definite terms.

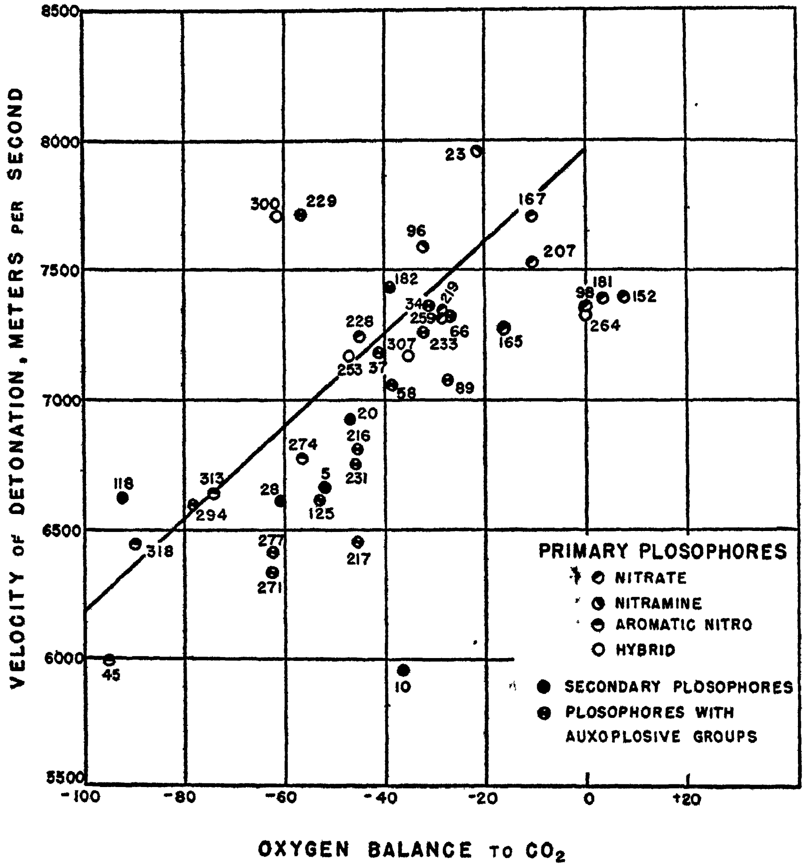


FIG. 13. Effect of oxygen balance on detonation velocity at a loading density of 1.50

V. CONSTITUTION AND HEAT OF EXPLOSION

The foregoing treatment of the subject, using plosophores and oxygen balance, has amounted to a rough system of thermodynamics. It is appropriate, therefore, to analyze the problem from that viewpoint and to reduce the qualitative theory to a quantitative basis. Such an approach is not original, having been developed essentially in its main outline during the last century (1). In the interest of completeness, however, it is briefly recapitulated here to permit a direct comparison with the organic approach given above. Of various energy terms which may be considered, the one known as the *heat of explosion* (H_E) must be the most significant, for by definition it is the energy liberated when the explosive compound is detonated to produce its detonation products. In practice this energy is difficult to measure directly, although it has been successfully accomplished by the use of a bomb calorimeter of adequate strength (19). Fortunately, it may be very closely approximated by a simple calculation from the heat of combustion (H_C) as determined by standard calorimetry methods (7, 18, 17, 24). In the following diagram the energy levels and the changes in energy in passing from one level to another are shown. The various H terms correspond to $-\Delta E$ in standard usage.

Explosive compound ($C_nH_mO_pN_q$)

H_C | H_E

Detonation products (C, CO, CO_2, H_2O, N_2)

H_C (products)

Complete combustion products (CO_2, H_2O, N_2, O_2)

From this it is clear that $H_E = H_C - H_{C(\text{products})}$ and that the last term should be easily obtained, since the values for the oxidation of the partially oxidized detonation products (hydrogen, carbon, carbon monoxide) are known with great accuracy.

This method of obtaining H_E requires that one know what the products of detonation are and the extent to which they are formed, and this in turn necessitates either an extremely difficult experimental technique which has yet to be devised or a reasonable assumption designed to allow a simple treatment of the problem. Fortunately the latter method in practice appears to be sufficient for all practical purposes.

In the light of the water gas reaction it can be assumed that, when an explosive detonates, the oxygen present combines first with the carbon to form carbon monoxide, then with the hydrogen to form water, and finally with the carbon monoxide to form carbon dioxide (9, 21, 25). The water formed is considered to be gaseous, as it is under the explosion conditions. In compounds of negative balance these reactions can be considered as occurring in steps and only so far as the oxygen will allow. In compounds of zero balance, the detonation products will coincide with the final combustion products, for then $H_{C(\text{products})}$ will equal zero and H_E will equal H_C.

There are grounds to doubt that the above assumption is correct to better than 15 per cent, and some investigators (2) have preferred an arrangement of reactions in which at least part of the hydrogen present is first converted to water, followed by oxidation of the carbon to carbon monoxide and then to carbon dioxide. Others have applied an arbitrary correction factor. The actual performance of specific explosives presumably lies at some indeterminate position between the extreme assumptions. In any case the differences due to the various assumptions will diminish with increasing oxygen balance and will disappear entirely at the point where oxygen is sufficient to convert all the hydrogen to water and all the carbon to carbon monoxide. This is the so-called carbon monoxide point and corresponds in oxygen balance to about −20 to −25 per cent. In the range of oxygen balance of greatest apparent interest, therefore, the difficulty does not exist.

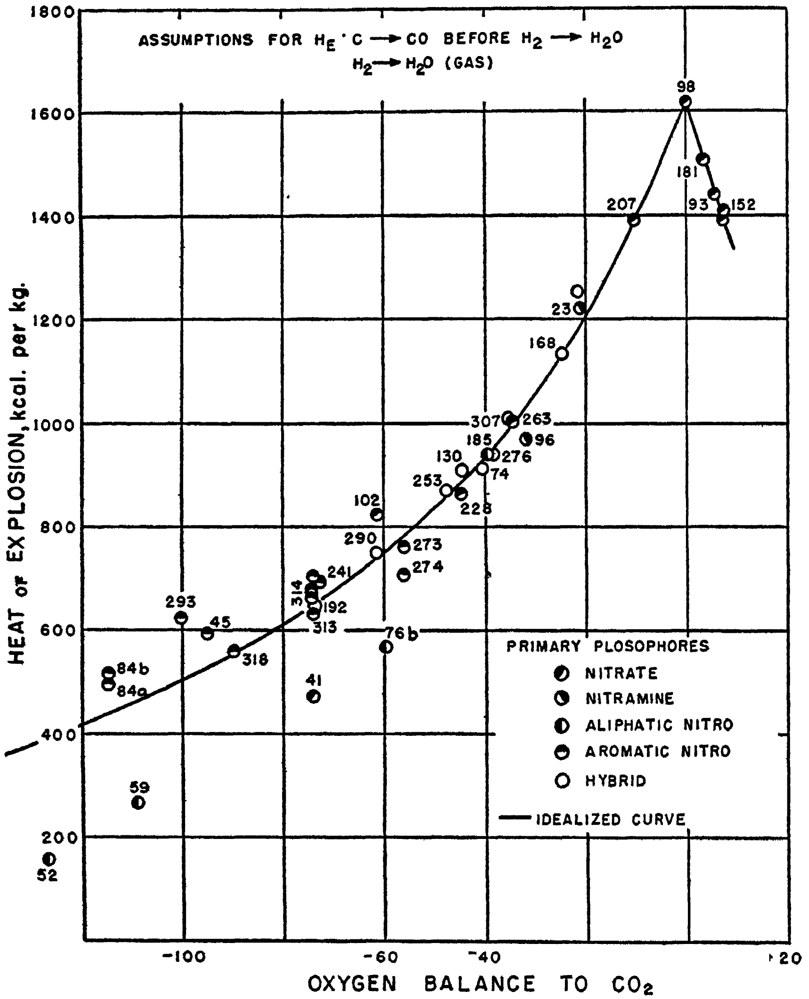


FIG. 14. Heat of explosion *vs.* oxygen balance for primary plosophores and their hybrids.

Using the first assumption with all explosives for which heat of combustion values have been available, H_E has been calculated on a weight basis and the resulting H_E values for primary plosophoric and hybrid molecules have been plotted against oxygen balance, yielding figure 14. It is seen that the heat of explosion varies directly with oxygen balance for these compounds, reaches a maximum value at zero balance (10), and in general resembles the power curves of figure 1. The results of a similar treatment of secondary plosophoric and auxoplosive compounds are given in figure 15. The poor showing of most of these in comparison with the idealized primary plosophoric curve is attributable to the partial combustion states that such groups represent and is the explanation for their generally unfavorable nature.

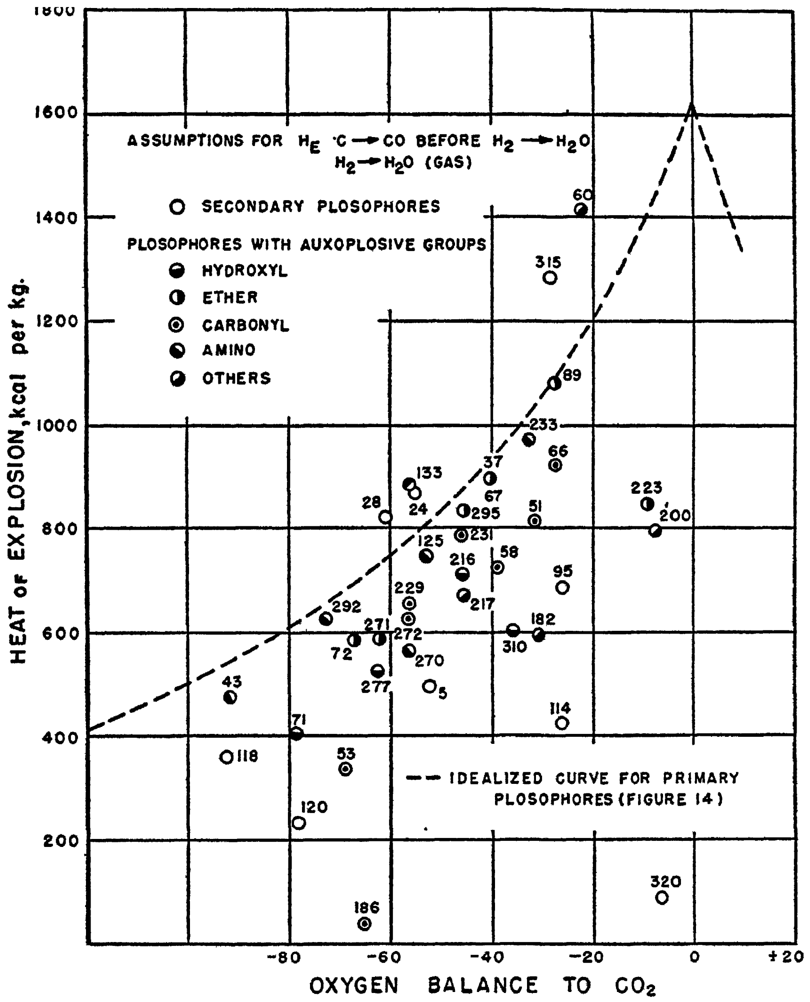


FIG. 15. Heat of explosion *vs.* oxygen balance for secondary plosophores and primary plosophores with auxoplosive groups.

The conclusions reached above from organic chemical considerations are explained and, as has been well known in the past, H_E is seen to be directly related to power (4). This last point is illustrated by figures 16 and 17, which present the relationship between lead block or ballistic mortar performance and heat of explosion together with idealized curves obtained by inspection.

With this point established, we have completed our objective of connecting the performance of explosives with their constitutions and with few exceptions have found the existing data consistent with the treatment developed. It remains, however, to extend the scope of the system to the prediction of power in advance of synthesis. By methods which are not novel (8, 21) but the details of which will not be presented here, it is possible to derive H_C values for each structural feature commonly present in primary plosophoric and hybrid explosives. In accord with standard thermodynamic reasoning, summation of the appropriate individual values for all the features of specified structure then yields the H_C value of the proposed explosive with an accuracy estimated to be ± 1 per cent and a maximum deviation of less than ± 3 per cent. From such a

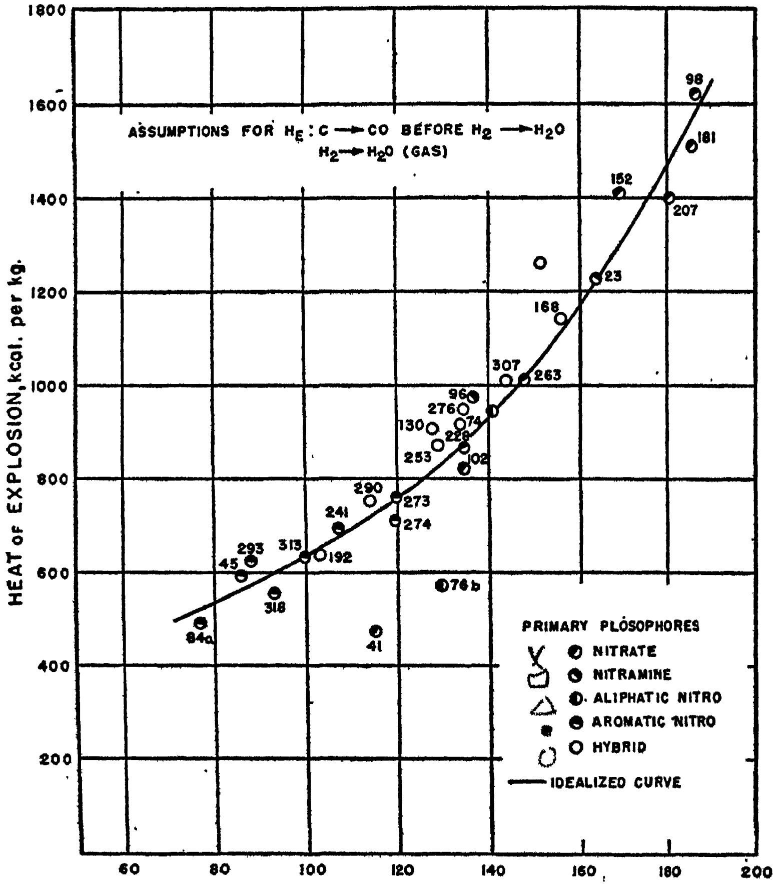


FIG. 16. Heat of explosion *vs.* power and brisance (lead block expansion test) of primary plosophores and their hybrids.

calculated value H_E may be obtained by the usual methods, and through figures 16 and 17 the expected power performance can then be estimated. Thus one has prior to synthesis a method of predicting performance of compounds containing the molecular structures reported.

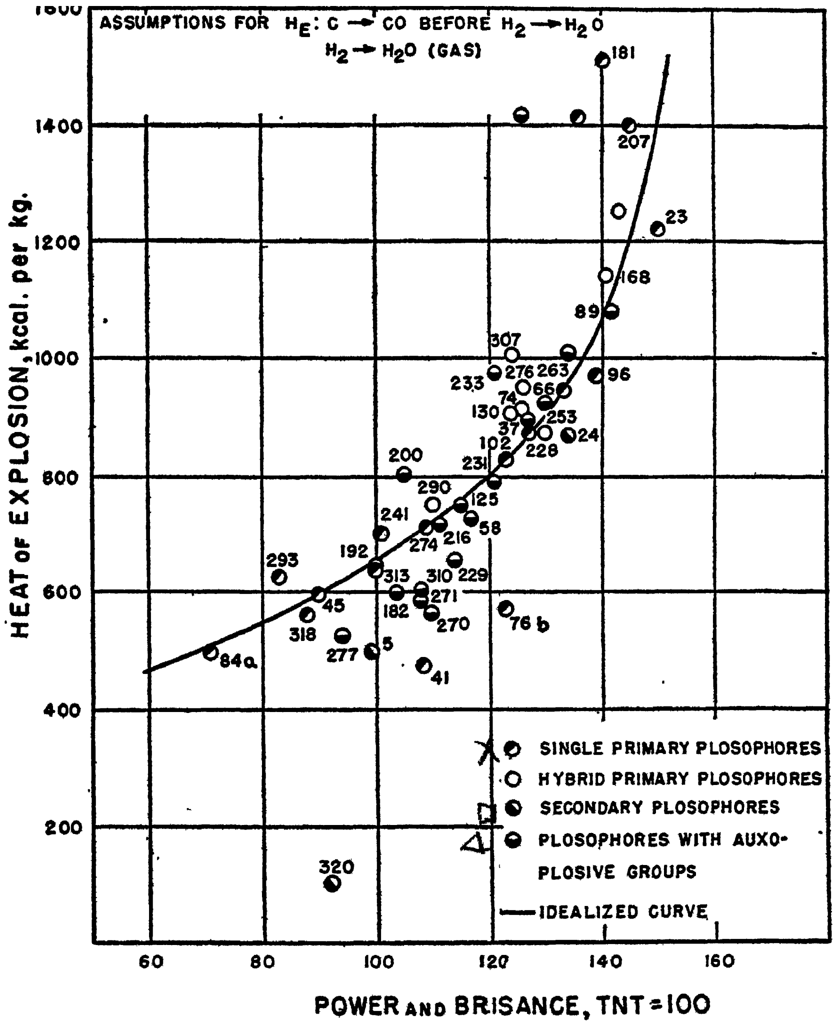


FIG. 17. Heat of explosion *vs.* power and brisance (ballistic mortar) of plosophoric compounds.

One might assume also that figures 1 to 5 or figure 10 would supply the power of a proposed structure directly from oxygen balance provided only primary plosophores are present. While use of these idealized curves will yield a reasonable prediction of performance for primary plosophoric and hybrid molecules, it is certain that the value which can be calculated will be a better approximation. One should probably not infer that either method is of such accuracy that further research looking toward the discovery of more powerful organic explosives is unwarranted, especially since the region of greatest interest (between −20 per cent and ± 0 per cent in oxygen balance) has been so little explored. Furthermore, the remarks made here are based entirely on presently available data, which, while intended to be complete, do not disclose fully the considerable advances in explosives theory made during the war nor contain examples of all appropriate groups which might be introduced into explosive molecules. There are also a number of assumptions which have entered into the thermodynamic treatment of the problem and whose net effect has not been accurately estimated. It is quite conceivable that future work will show that certain new arrangements of atoms may give enhanced performance beyond that which would be expected from oxygen balance or from *calculated* heat of combustion.

VI. SUMMARY

Much available data on the power of pure organic explosives have been assembled and interpreted. Using the terms oxygen balance, plosophoric and auxoplosive groups, and heat of explosion, the types of molecules have been defined which appear of synthetic interest for use as high explosives. It is hoped that the new qualitative treatment of the subject will prove a stimulation to renewed synthetic organic research in a difficult field.

VII. REFERENCES

(1) Berthelot, M. P. E.: *Explosives and their Power*, translated by C. N. Hake and W. MacNab. Murray, London (1892).
(2) Brinkley, S. R., and Wilson, E. B., Jr.: OSRD Report No. 905, September, 1942.
(3) Brown, F. W.: U. S. Bur. Mines, Tech. Paper **643**, 26 pp. (1942).
(4) Brunswig, H.: *Explosives*, translated by C. E. Munroe and A. L. Kibler. John Wiley and Sons, Inc., New York (1912).
(5) Davis, T. L.: *The Chemistry of Powder and Explosives*, Vol. I, Vol. II. John Wiley and Sons, Inc., New York (1941, 1943).
(6) Duin, C. F. van, and Lennep, B. C. R. van: Rec. trav. chim. **39**, 145 (1920).
(7) Garner, W. E., and Abernethy, C. L.: Proc. Roy. Soc. (London) **A99**, 213 (1921).
(8) Kharasch, M. S.: Bur. Standards J. Research **2**, 359 (1929).
(9) Kistiakowsky, G. B., and Wilson, E. B., Jr.: OSRD Report No. 69, January, 1941.
(10) Marshall, A.: *Explosives*, 2nd edition, Vol. I. J. and A. Churchill, London (1917).
(11) Moser, L.: Oesterr. Chem.-Ztg. **20**, 193 (1917).
(12) Munroe, C. E., and Tiffany, J. E.: U. S. Bur. Mines, Bull. **346**, 148 pp. (1931).
(13) Naoum, P.: *Nitroglycerine and Nitroglycerine Explosives*, translated by E. M. Symmes. The Williams & Wilkins Company, Baltimore, Maryland (1928).
(14) Perrott, G. St. J., Tolch, N. A., and Crawshaw, J. E.: Am. Inst. Mining Met. Engrs., Tech. Pub. **101**, 10 pp. (1928).
(15) Pletz, V.: J. Gen. Chem. (U.S.S.R.) **5**, 173 (1935).
(16) Proceedings of the International Congress for Applied Chemistry, Vol. II, p. 463, Berlin (1904).
(17) Rinkenbach, W. H.: J. Am. Chem. Soc. **52**, 116 (1930).
(18) Robertson, R.: J. Chem. Soc. **119**, 1 (1921).
(19) Robertson, R., and Garner, W. E.: Proc. Roy. Soc. (London) **A103**, 539 (1923).
(20) Roth, J. F.: Z. ges. Schiess-u. Sprengstoffw. **36**, 4, 28, 52 (1941).
(21) Schmidt, A.: Z. ges. Schiess-u. Sprengstoffw. **29**, 259, 296 (1934); **30**, 364 (1935).

(22) Stettbacher, A.: Z. ges. Schiess-u. Sprengstoffw. **13**, 225 (1918); **14**, 220 (1919); **15**, 165 (1920); **25**, 439 (1930); **37**, 42, 62 (1942); Die Schiess-und Sprengstoffe, 2nd edition, J. A. Barth, Leipzig (1933).

(23) Storm, C. G., and Cope, W. C.: U. S. Bur. Mines, Tech. Paper **125**, 64 pp. (1916).

(24) Taylor, J., and Hall, C. R. L.: J. Phys. Colloid Chem. **51**, 593 (1947).

(25) Taylor, J., Hall, C. R. L., and Thomas, H.: J. Phys. Colloid Chem. **51**, 580 (1947).

(26) Taylor, W., and Morris, G.: Trans. Faraday Soc. **28**, 545 (1932).

(27) Tolch, N. A., and Perrott, G. St. J.: U. S. Bur. Mines, Repts. Investigations **3039**, 10 pp. (1930).

THE CHEMISTRY OF THE OXAZOLINES

RICHARD H. WILEY AND LEONARD L. BENNETT, JR.

Department of Chemistry, University of North Carolina, Chapel Hill, North Carolina

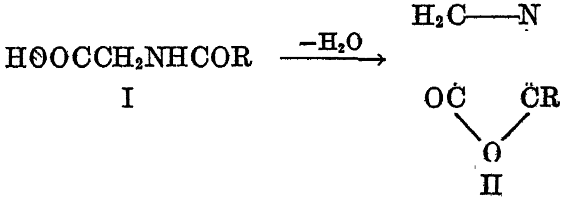
Received April 3, 1948

CONTENTS

I. Introduction 447
II. 2-Oxazolines 449
  A. Syntheses of 2-oxazolines 449
    1. From β-haloalkylamides 449
      Table 1. Oxazolines from β-haloalkylamides 451
    2. From β-hydroxyalkylamides 452
      Table 2. Oxazolines from β-hydroxyalkylamides 452
    3. From amides of β,γ-unsaturated amines 455
    4. From iminoesters 455
  B. Physical properties of 2-oxazolines 456
    Table 3. Boiling and melting points of 2-oxazolines 456
  C. Chemical properties of 2-oxazolines 457
    1. Salt formation 457
    2. Hydrolysis 457
    3. Reduction 459
    4. Oxidation 460
    5. The iminoester–oxazoline–amide equilibrium 460
    6. Activity of 2-methyl group 461
    7. Synthetic applications 461
  D. Commercial and industrial uses of 2-oxazolines 462
III. Substituted 2-oxazolines 463
  A. Pseudoureas (2-amino-2-oxazolines) 464
    1. Syntheses of pseudoureas 464
      (a) From β-haloalkylureas 464
      (b) From β-hydroxyalkylureas 465
      (c) From sodium cyanamide and chlorohydrins 466
        Table 4. Pseudoureas: 2-amino-2-oxazolines 466
    2. Reactions of pseudoureas 467
  B. 2-Mercapto-2-oxazolines 469
  C. 4-Keto-2-oxazolines 470
  D. 5-Keto-2-oxazolines 471
    Table 5. 4-Substituted-5-keto-2-oxazolines 472
IV. 3-Oxazolines 472
V. 4-Oxazolines 473
  Table 6. 4-Oxazolines 473
VI. References 474

I. INTRODUCTION

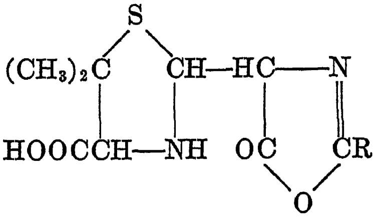
The oxazolines, or dihydroöxazoles, are of interest because of the close relation between the acyl derivatives of α-amino acids (I) and the 5-keto-2-oxazolines (II).

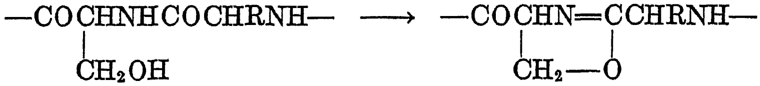


I

II

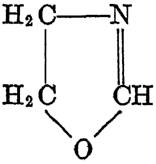
The Erlenmeyer azlactone synthesis of amino acids (21), the formal possibility of the existence of oxazoline rings in polypeptides (7, 89), the azlactone penicillin possibility (III) (24a, 81a), and the use of azlactones in building up chains with peptide linkages (6a) are all manifestations of this relation. Other interests in the oxazolines have also been developed. The formation of an oxazoline ring through cyclodehydration at peptide linkages formed from the α-amino-β-hydroxy acids, serine or threonine, has been postulated by Blackburn and others (13).

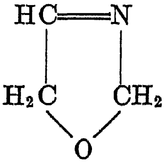

III

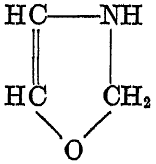

$$-CO\underset{\displaystyle CH_2OH}{CH}NHCOCHRNH- \longrightarrow -CO\underset{\displaystyle CH_2-O}{CHN=C}CHRNH-$$

The 2-oxazolines, available from *N*-acyl derivatives of ethanolamine, have been used as cationic surface-active agents Analogies drawn from the relation between procaine and 2-oxazolines have been used in designing local anesthetics (1, 53). One oxazoline has been reportedly isolated from a Canadian weed (42). Synthetic anhydrides of acylserines (7) and other peptides (40, 76) have been studied.

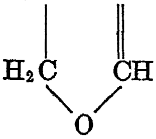
There are three possible types of oxazoline or dihydroöxazole: 2-oxazoline (IV), 3-oxazoline (V), and 4-oxazoline (VI).

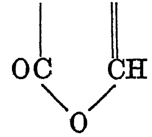

IV

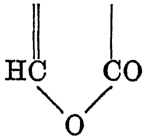

V


VI

2-Oxazolines (IV) and their derivatives are by far the most common. Very little is known about the 3- and 4-oxazolines. Ketoöxazolines (VII, VIII, IX), in which a carbonyl replaces a methylene group, derived from 2- and 4-oxazolines are known, but those from 3-oxazolines are not known.[1] The 5-keto-2-oxazolines


VII

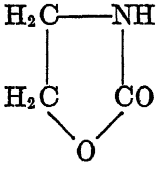

VIII

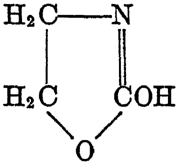

IX

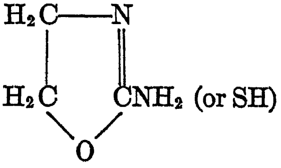
---

[1] *Note added in proof:* Compounds of this type are reported in *The Chemistry of Penicillin*, Princeton University Press, 1949, pp. 739–42, where they are called pseudoxazolones.

(VIII) or azlactones, the intermediates in the Erlenmeyer azlactone synthesis of amino acids, have been reviewed by Carter (21) and only recent developments will be included here.[2] Some oxazolidones, where the carbonyl is on a carbon adjacent to an NH group as in X, are hydroxyoxazolines in their enolic forms (XI). These belong properly with a discussion of other oxazolidones and are not included here. The 2-amino(or mercapto)-2-oxazolines (XII) are included in separate sections of this discussion. The oxidoöxazoles (which are 4,5-epoxy-oxazolines) and the 2-aminoöxazoles (or 2-imino-2-oxazolines) have been discussed elsewhere (89).


X


XI


XII

Historically, the first oxazoline was prepared in 1884 by Andreasch (2). He suspected that the compound resulting from the dehydrohalogenation of allylurea bromide contained a new cyclic structure, but failed to deduce its correct formula. Gabriel five years later (32) first recognized an oxazoline as such and began an extensive study of the chemistry of this heterocyclic system. The present review has been prepared to analyze the available information on the oxazoline types and to indicate portions of the field for which little or no information is available.

II. 2-Oxazolines

A. SYNTHESES OF 2-OXAZOLINES

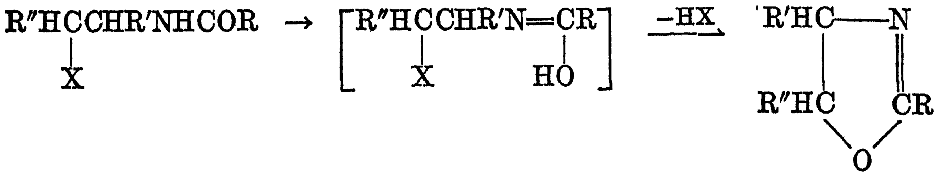
The known syntheses for 2-oxazolines follow the usual pattern for the formation of cyclic compounds. The ring is formed by a condensation reaction in which the elements of water or halogen acid are removed from a suitable open-chain compound. Examples of the types used are *N*-acyl derivatives of β-halo- and β-hydroxy-alkylamines and iminoesters of β-halo- or β-hydroxy-amines. A few types of substituted 2-oxazolines are also available by cyclization reactions from modified amides. Thus, β-halo- and β-hydroxy-alkylureas form 2-amino-2-oxazolines, and acylamino acids form the azlactones or 5-keto-2-oxazolines. 2-Mercapto types are available directly from carbon bisulfide and ethanolamines. These, and other syntheses, are discussed in this and following sections.

1. From β-haloalkylamides

The synthesis of 2-oxazolines from β-haloalkylamides has been the most widely used method for preparing these compounds. The reaction takes place according to the following equation:

---

[2] *Note added in proof:* These compounds are also surveyed in *The Chemistry of Penicillin*, Princeton University Press, 1949.



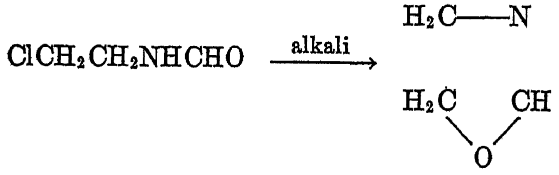
XIII

The various 2-oxazolines prepared by this method are listed in table 1. R in XIII, which is derived from the acyl group, has been varied widely but in most reported 2-oxazolines is aryl. Variations in R′ and R″ in XIII are very few, since only a few β-haloamines are available. In all but a few 2-oxazolines, R′ is H and R″ is H or CH_3. These are derived from β-haloethyl- or β-halopropyl-amines. The former is readily available from ethanolamine; the latter from allylamine and hydrogen bromide (25) or from nitrosoacetone as follows (73):

$$CH_3COCH_2NO \rightarrow CH_3CHOHCH_2NH_2 \xrightarrow{HBr} CH_3CHBrCH_2NH_2\cdot HBr$$

β-Chloro-*n*-butylamine, from which a 2-oxazoline with R″ an ethyl group is obtained, has been prepared by Bookman (15). Since the haloamines are usually prepared from hydroxyamines, which can themselves be converted to oxazolines, the haloamide method is not as direct as is the synthesis from hydroxyamides described in the next section.

In most instances the dehydrohalogenation has been effected by heating the β-haloamide with aqueous or alcoholic alkali; acetic anhydride has been used, but yields are low (33, 36). The oxazoline is easily isolated and purified. A typical preparation is that used by Gabriel and Heymann (36) for 2-phenyl-2-oxazoline. One mole of bromoethylbenzamide is dissolved as quickly as possible in hot water and 1 mole of sodium hydroxide added; an oil separates which is steam distilled, extracted with ether, and purified by distillation. It was this method that Wenker (86) used in preparing unsubstituted 2-oxazoline from β-chloroethylformamide, and 2,2′-Δ^2-dioxazoline from di-β-chloroethyloxamide.



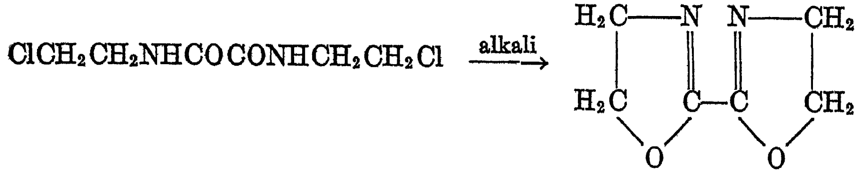
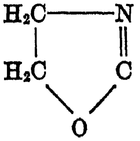


TABLE 1

2-Oxazolines from β-haloalkylamides

SUBSTITUENTS IN 2-OXAZOLINE*			YIELD†	REFERENCE
R	R′	R″		
			per cent	
H	H	H	39	(86)
CH_3	H	H	40	(33, 35, 36)
C_6H_5	H	H	50	(36)
p-$CH_3C_6H_4$	H	H	75	(66)
o-$CH_3C_6H_4$	H	H	50	(66)
$C_6H_5CH_2$	H	H	Poor	(25)
m-$NO_2C_6H_4$	H	H	82	(25, 53)
o-$NO_2C_6H_4$	H	H	92	(53)
p-$NO_2C_6H_4$	H	H	72	(53)
p-$ClC_6H_4CH_2$	H	H	93	(53)
β-$C_{10}H_7$	H	H		(67)
α-$C_{10}H_7$	H	H		(67)
p-$CH_3OC_6H_4$	H	H		(62)
p-$C_2H_5OC_6H_4$	H	H		(26)
p-$C_4H_9OC_6H_4$	H	H		(26)
m-NO_2-p-$CH_3OC_6H_3$	H	H	91	(53)
m-$NO_2C_6H_4CH{=}CH$	H	H	75	(53)
	H	H	84	(86)
C_6H_5	C_6H_5	H		(34)
p-$NO_2C_6H_4$	C_4H_9	H	43	(53)
p-$NO_2C_6H_4$	C_6H_5	H	86	(53)
CH_3	H	CH_3	15	(79)
C_6H_5	H	CH_3	60–66	(36, 79)
C_6H_5	H	C_2H_5		(15)
$C_6H_5CH_2$	H	CH_3	Poor	(25)
$C_6H_5CH{=}CH$	H	CH_3	60	(25)
C_6H_5	H	C_6H_5		(91)
o-$C_6H_4CH_3$	H	CH_3	65	(66)
p-$C_6H_4CH_3$	H	CH_3		(66)
o-$C_6H_4NO_2$	H	CH_3	83	(53, 79)
p-$C_6H_4NO_2$	H	CH_3	91	(53, 79)
m-$C_6H_4NO_2$	H	CH_3	93	(25, 53)
α-$C_{10}H_7$	H	CH_3		(67)
β-$C_{10}H_7$	H	CH_3		(67)
C_6H_5	CH_3	CH_3		(73)
p-$C_6H_4NO_2$	CH_3	CH_3	74	(53)
C_6H_5	CH_3	C_2H_5		(44)

* R, R′, R″ of formula XIII.

† Where there is more than one reference, only the best yield is given.

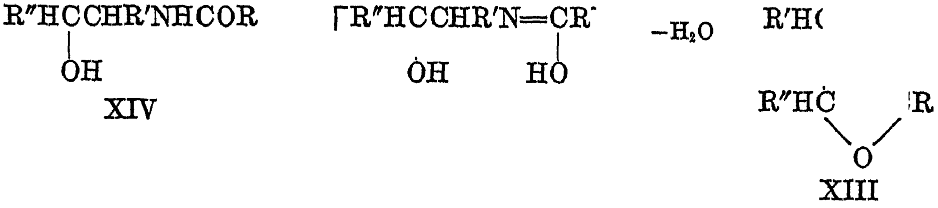
TABLE 2

Oxazolines from β-hydroxyalkylamides

SUBSTITUENTS IN 2-OXAZOLINE			YIELD	REFERENCE
In 2-position	In 4-position	In 5-position	per cent	
CH_3	H	H	30	(85)
C_2H_5	H	H	35	(85)
C_6H_5	H	H	22	(85)
$CH_3(CH_2)_{10}$	H	$ClCH_2$		(8)
C_6H_5	H	C_6H_5		(74)
C_6H_5	H	$C_6H_5COOCH_2$		(4)
C_6H_5	H	$ClCH_2$		(3)
C_6H_5	H	$ClCOCH_2$		(6)
C_6H_5	H	$C_2H_5OCOCH_2$		(6)
p-$NO_2C_6H_4$	H	$ClCH_2$	88	(53)
p-$NO_2C_6H_4$	H	$(C_2H_5)_2NCH_2$	86	(53)
p-$NO_2C_6H_4$	H	$(C_4H_9)_2NCH_2$	97	(53)
C_6H_5	H	CH_3, CH_3		(22)
C_6H_5	H	C_6H_5, CH_3	95	(51)
m-$NO_2C_6H_4$	H	CH_3, CH_3	95	(53)
p-$NO_2C_6H_4$	H	CH_3, CH_3	86	(53)
C_6H_5	CH_3OCO	H		(7)
CH_3	CH_3, CH_3	H		(77)
CH_3	$HOCH_2$, $HOCH_2$	H		(79a)
CH_3	CH_3, CH_3COOCH_2	H		(77)
CH_3	CH_3COOCH_2, CH_3COOCH_2	H		(77, 78)
$C_{11}H_{23}$	CH_3, CH_3	H		(77)
C_6H_5	CH_3, $HOCH_2$	H	67	(10)
C_6H_5	C_2H_5, $HOCH_2$	H	72	(10)
C_6H_5	$HOCH_2$, $HOCH_2$	H	60-65	(10)
C_6H_5	CH_3, CH_3	H		(77)
C_6H_5	CH_3	C_6H_5	75	(61)
CH_3	C_6H_5	C_6H_5, C_6H_5		(51)
C_6H_5	CH_3	$C_6H_5CH_2$, $C_6H_5CH_2$		(9)
p-$NO_2C_6H_4$	(4,5-Cyclohexano)		38	(53)

2. *From β-hydroxyalkylamides*

The preparation of 2-oxazolines by the dehydration of β-hydroxyalkylamides is similar to the preparation from β-haloalkylamides. The reaction can be


XIV

XIII

visualized as a cyclodehydration of the imido form of an amide of a β-hydroxyalkylamine. The 2-oxazolines prepared by this method are listed in table 2. β-Hydroxyamides are available from amines prepared in turn from chlorohydrins on reaction with ammonia; or via the Gabriel primary amine synthesis (22); or from the condensation of aromatic aldehydes with nitroethane (61); or by reduction of cyanohydrins (91). Many types are available and can be used in preparing 2-oxazolines with a variety of substituents. Several examples of oxazolines with two substituents in the 4- or 5-position are known and many types with functional groups as substituents have been prepared. Thus, the glycerol chlorohydrins react with amines to furnish hydroxyamides for synthesis of the hydroxymethyl-substituted types. This method is more direct than the haloamide method, discussed in the preceding section, because the halo derivatives necessary for the preparation of haloamides are usually prepared from hydroxyamines. This probably accounts for the wide variety of 2-oxazolines prepared from hydroxyamides.

Among the dehydrating agents, sulfuric acid (51, 53, 61, 74) and thionyl chloride (4, 7, 53) have been most commonly used; phosphorus oxychloride (4) and phosphorus pentoxide (51, 85) are also effective. In one instance (22) cyclization occurred upon heating with water and cyclization may take place on heating in organic solvents. Thus, Valco (79a) dehydrated the hydroxyamide by boiling in toluene or xylene and distilling out the water–hydrocarbon azeotrope as the water formed.

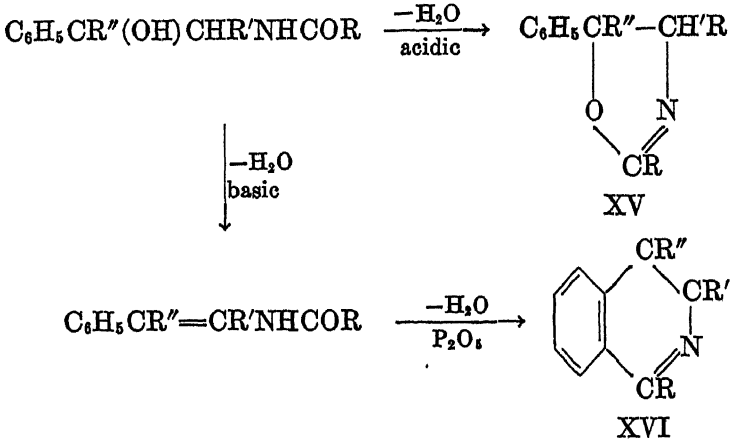
The mechanism is probably cyclodehydration as indicated above. It might be supposed, however, to involve preliminary dehydration followed by cyclization:

$$\text{XIX} \rightarrow [\text{R}''\text{HC}{=}\text{CR}'\text{NHCOR}] \rightarrow \text{XIII}$$

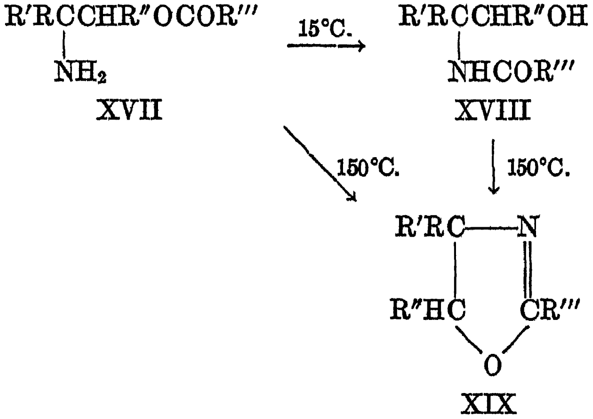
This point has not been systematically investigated and there is not sufficient evidence to indicate the course taken by the reaction. Information on this point might be obtained by comparing the yields of oxazoline obtained from unsaturated amides with yields obtained from β-halo- and β-hydroxy-amides. The work of Krabbe *et al.* (51) discussed below suggests that unsaturated amides are not readily converted to oxazolines.

The isolation and purification of the product are simple and straightforward. Typical procedures are those of Leffler and Adams (53), using either thionyl chloride or sulfuric acid. The procedure for the sulfuric acid dehydration is as follows: One hundred cubic centimeters of concentrated sulfuric acid is added to 0.15 mole of β-hydroxy-*p*-nitrobenzamide and the mixture held at 55–60°C. for 10 min. After cooling to 15°C. it is poured into ice water and filtered, and the oxazoline precipitated from the filtrate by adding ammonia. The procedure for the thionyl chloride dehydration is described as follows. Thionyl chloride (33.4 g.) is added to 0.01 mole of γ-diethylamino-β-hydroxypropyl-*p*-nitrobenzamide and the mixture heated on the water bath for 1.5–2 hr. After cooling to 5°C., it is poured into dry ether and allowed to stand overnight. The hydrochloride of the oxazoline separates and is collected on a filter and dissolved in water. The free oxazoline is obtained by making the filtrate alkaline with sodium hydroxide.

In certain β-hydroxyamides containing a β-phenyl substituent there are two possibilities for ring closure to give either the isoquinoline or the oxazoline, depending upon the manner in which the dehydration is carried out. Krabbe *et al.* (51) have found that acidic reagents, such as sulfuric acid, phosphorus pentoxide, or chloroxalic ester, gave the oxazoline (XV), while ethylmagnesium bromide gave exclusively the *N*-acylvinylamine, which could then be converted to the isoquinoline (XVI) with phosphorus pentoxide. Where R″ is phenyl, only the isoquinoline can be formed.



A variation of this method of synthesis uses β-aminoesters (XVII) as starting materials (49, 77). The transformation probably involves migration of the acyl group from oxygen to nitrogen to yield XVIII prior to cyclization to XIX. Such migrations are well-known and will be discussed in a later section.

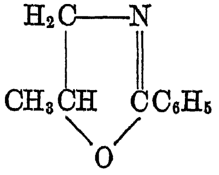


If an amide of the β-aminoester (XVII) is used in place of the aminoester, an acid is split out with the formation of the oxazoline. When R‴ is CH_3, R″ is H, and R and R′ are acetoxymethyl groups, 2-methyl-4,4-bis(acetoxymethyl)-2-oxazoline is formed in 67 per cent yield. (78). A series of oxazolines have been prepared from *N*-acyl derivatives of amino sugars by reactions which fall within this class (16, 88, 92, 93).

3. *From amides of unsaturated amines*

Gabriel and Stelzner (38) observed that *N*-vinylbenzamide was converted to 2-phenyl-2-oxazoline on distillation. This reaction probably involves simple cyclization by an addition reaction. It has also been observed, by Kay (50) and Diels and Beccard (23), that *N*-allylbenzamide is converted to 2-phenyl-5-methyl-2-oxazoline (XX) on heating with concentrated sulfuric or hydrochloric acid. The formation of this product can be explained on the assumption that an intermediate β-hydroxyalkylamide sulfate or β-haloalkylamide is formed by addition of acid to the allyl double bond, followed by cyclization.

$$CH_2{=}CHCH_2NHCOC_6H_5 \rightarrow [CH_3CH(OSO_2OH)CH_2NHCOC_6H_5] \rightarrow$$

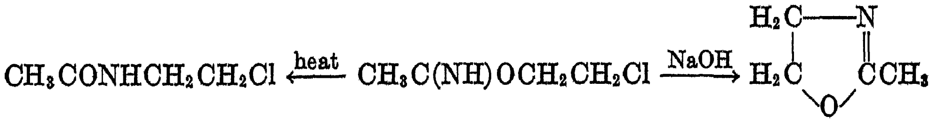


XX

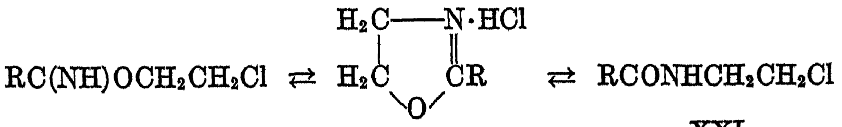
2-*o*-Hydroxyphenyl-5-methyl-2-oxazoline was apparently prepared by this method (23). The type reaction has not been otherwise used, presumably because the unsaturated amides have been available only from β-hydroxy- or β-halo-alkylamides which are themselves readily cyclized to 2-oxazolines. The work of Krabbe *et al.* (51), discussed previously, indicates that *N*-acylvinylamines having two phenyl groups on the β-carbon atom will cyclize to isoquinolines (XVI).

4. *From iminoesters*

Another route to oxazolines starts with the iminoesters of β-halo- or β-hydroxyamines. Gabriel and Neumann (37) showed that the oxazoline is formed when the iminoester of a β-chloroalcohol is decomposed with sodium hydroxide or in a desiccator over sulfuric acid, but that the *N*-chloroalkylamide product forms when the iminoester is decomposed thermally.



Wislicenus and Körber (90) investigated this conversion more thoroughly and were able to prove that the oxazoline is the intermediate in the migration of an alkyl group from the oxygen in an iminoester to the nitrogen in an amide (XXI).



XXI

This equilibrium will be discussed in detail in the section on reactions of 2-oxazolines.

The iminoester method has not been extensively used. A recent modification (54) yields the oxazoline in one step from a nitrile and a β-aminoalcohol. The nitrile and alcohol, with a little sodium methoxide as catalyst, are heated together in a silver-lined vessel until evolution of ammonia has ceased, and the oxazoline is then distilled at reduced pressure from the reaction mixture. 2-[(Ethoxymethoxy)methyl]-2-oxazoline, prepared by this method from ethanolamine and $C_2H_5OCH_2OCH_2CN$, was obtained in 32 per cent yield.

Another variation (14) starts with a β-hydroxyamine and an iminoester. The first step is presumably the formation of a new iminoester, which then cyclizes normally. This method was used by Bockemühl and Knoll (14) in the following synthesis in which the reaction was carried out in absolute alcohol, the mixture allowed to stand, and the oxazoline obtained by evaporation of the solvent.

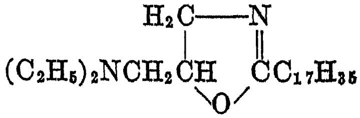
TABLE 3

Boiling and melting points of 2-oxazolines

SUBSTITUENTS IN 2-OXAZOLINE	BOILING POINT	MELTING POINT	REFERENCE
	°C.	°C.	
	98		(86)
2-Methyl	109–110/757 mm.		(35)
2-Phenyl	242–243		(36)
2-*p*-Tolyl	264–265	66	(66)
2-*o*-Tolyl	254		(66)
2,5-Dimethyl	117–119		(79)
2-Phenyl-5-methyl	243–244		(36)
2,5-Diphenyl	229–230/44 mm.		(91)
2,4-Diphenyl	210/19 mm.		(34)

$$(C_2H_5)_2NCH_2CHOHCH_2NH_2 + C_{17}H_{35}C(NH)OC_2H_5 \xrightarrow{-NH_3}$$

$$(C_2H_5)_2NCH_2CHOHCH_2N{=}C(OC_2H_5)C_{17}H_{35} \xrightarrow{-C_2H_5OH}$$



B. PHYSICAL PROPERTIES OF 2-OXAZOLINES

The physical properties of the oxazolines depend significantly upon the nature of the substituents in the ring. The low-molecular-weight members are mobile colorless oils with a sweet, pyridine-like odor. 2-Oxazoline itself boils at 98°C. The effect of introducing substituents of increasing size is shown in table 3. Successful distillation at the boiling points given indicates a fair degree of thermal stability. Spectrochemical data have been reported for 2-phenyl-2-oxazoline (18) and 2,5-diphenyl-2-oxazoline (82).

C. CHEMICAL PROPERTIES OF 2-OXAZOLINES

The oxazoline ring itself imparts no unusual identity to the oxazolines as a class. They react, in general, about as one would predict that similar functioning groups in open-chain compounds would react. By way of summary it can be said that the 2-oxazolines are typical weak bases, whose lower members dissolve in water to render it alkaline to phenolphthalein. As bases they form the usual salts with acids; these salts vary in stability with the nature of the oxazoline. Quaternary salts are also known. The oxazoline ring is stable towards mild oxidation–reduction or hydrolytic conditions. 2-Phenyl-2-oxazolines can be nitrated and the nitro compound reduced to the amine without rupturing the ring (1, 53). Oxazolines containing a methylol group in the 4-position have been oxidized to oxazoline-4-carboxylic acids (10). Under vigorous conditions, however, the ring readily cleaves. Sodium and alcohol (39) yield a secondary amine. Heating with dilute acid, but apparently not always with dilute alkali, hydrolyzes oxazolines completely. It is noted that no instance of reduction of an oxazoline to an oxazolidine or dehydrogenation to an oxazole has been recorded. Other reactions, discussed in the following sections, provide additional evidence as to the behavior of the functioning groups in the 2-oxazolines.

1. Salt formation

Oxazoline salts of most of the common acid reagents have been reported. The stability of these salts is varied. 2-Oxazoline forms an unstable hydrochloride which spontaneously hydrolyzes to *N*-β-bromoethylformamide, but hydrochlorides of oxazolines of higher molecular weight are stable substances with definite melting points. Many salts have been used for identification of the oxazolines. The salts are of course unstable towards hot water, owing to the ease with which dilute acid hydrolyzes the ring. The properties of these salts are in general similar to those of salts of weak bases.

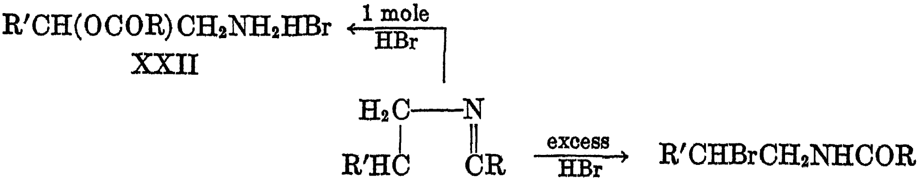
Only recently and only in two instances (41, 43) has the formation of quaternary oxazoline salts been recorded. Hamer and Rathbone (41) have made the methiodide and ethiodide of 2-methyl-2-oxazoline. The oxazoline reacts exothermally with methyl iodide to produce the quaternary salt as a white deliquescent solid; the corresponding ethiodide, a yellow deliquescent solid, is formed when the oxazoline is heated with ethyl iodide at 100°C. for 4 hr.

2. Hydrolysis

Oxazolines are not easily hydrolyzed by water, as is evident from the fact that the 2-aryl types are not decomposed when isolated by steam distillation. The importance of the role played by substituents in the nucleus is seen in the fact that Wenker reported (86) that boiling water converts 2-oxazoline to β-hydroxyethylformamide.

From the method employed for their synthesis, it is evident that oxazolines as a group have considerable stability towards aqueous or alcoholic alkali and such acidic dehydrating agents as thionyl chloride, sulfuric acid, and phosphorus pentoxide. Dilute aqueous acid has no effect on most oxazolines other than salt formation, but boiling with dilute acid gives an easy hydrolysis. Aqueous alkali can also rupture the ring, although apparently less readily than acid.

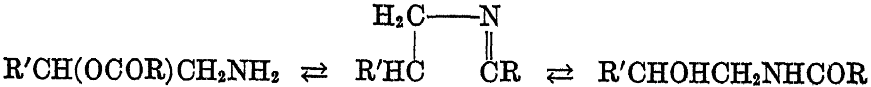
Gabriel and Heymann (36) early observed that an oxazoline might be hydrolyzed to a β-aminoester (XXII) with 1 mole of hydrobromic acid or to a β-haloalkylamide (XXIII) with an excess of hydrobromic acid. This reaction was carried out with 2-phenyl-2-oxazoline and 2-phenyl-5-methyl-2-oxazoline. Salomon (66) obtained similar results with 2-o-tolyl-2-oxazoline.



XXII

XXIII

Connected with the hydrolysis of oxazolines is the migration of an acyl group from the oxygen of a β-aminoester (XXIV) to the nitrogen of a β-hydroxyamide (XXV) and the reverse reaction. This migration has been frequently observed, and it has been definitely established that the ozaxoline is the intermediate in the transformation.



XXIV

XXV

The ease of this migration and the reagents required to effect it depend upon the nature of the compounds. Bettzieche (9), working with benzamides (XXVa)

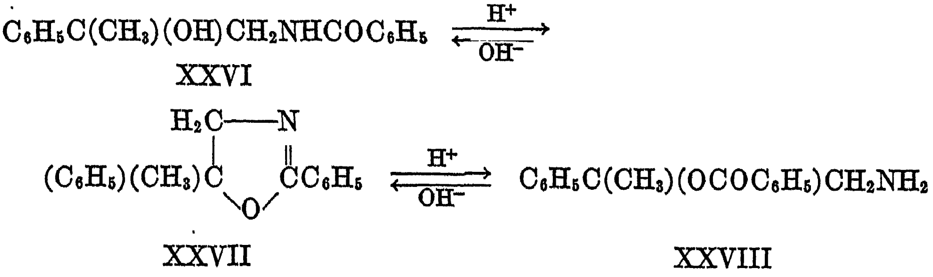
$$R'_2COHCH_2NHCOC_6H_5$$

XXVa

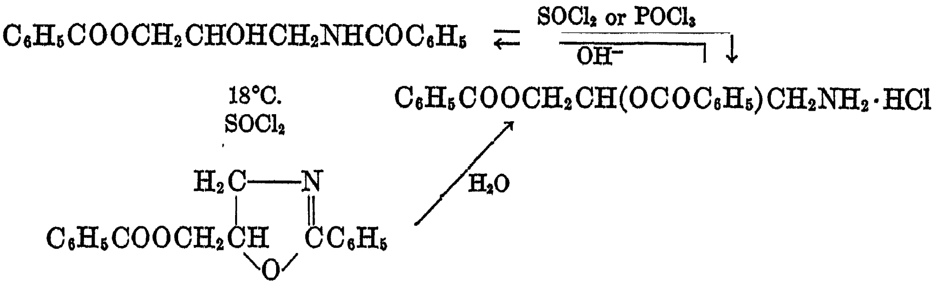
found that sulfuric acid rapidly caused migration from nitrogen to oxygen and quantitatively when R′ is methyl or ethyl. When R′ is benzyl or phenyl, migration was not observed. He also found that migration from oxygen to nitrogen under the influence of alkali was slow and gave mixtures of oxazolines and *N*-acyl compounds. On the other hand, Bergmann and Miekeley (7), in a study of benzoylserine, found that the cycle

N-benzoyl → oxazoline → *O*-benzoyl → *N*-benzoyl

could be completed without difficulty. In their studies of amides such as XXVI, Krabbe, Eisenlohr, and Schöne (51) found that they could move from left to right with acid and from right to left with base as indicated. If the oxazoline XXVII is isolated, it can be converted to the *N*-acyl compound (XXVI) with alkali or to the *O*-acyl compound (XXVIII) with acid.



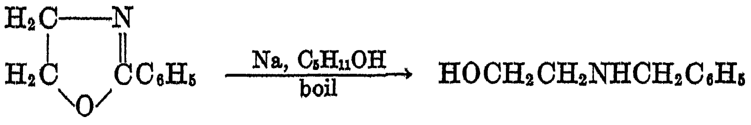
Similar results were obtained by Bergmann and Brand (4), who observed the migrations indicated in the following scheme:



It seems rather well established that the migration is from nitrogen to oxygen in acid and the reverse in alkali. That there may be some misunderstood or variable factors is indicated by the fact, noted in the formation of XXII and XXIII, that some oxazolines can be converted to either the *N*-acyl or *O*-acyl compounds by using hydrobromic acid in varying proportions. A fact of interest in connection with the hydrolytic stability of 2-oxazolines to acids is a ring cleavage in non-aqueous medium. 2,5-Diphenyloxazoline when dissolved in dry ether and treated with hydrogen chloride forms $C_6H_5CHClCH_2NHCOC_6H_5$ (74).

3. Reduction

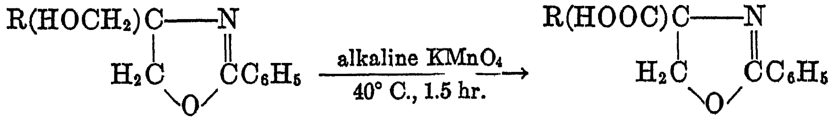
Attempts to reduce 2-oxazolines to oxazolidines using sodium and alcohol lead only to ring cleavage. Thus, with 2-phenyl-2-oxazoline there results *N*-benzylethanolamine (39).



No less vigorous reducing agents have been tried and this information does not indicate whether the 2-oxazoline or the oxazolidine ring is cleaved in the process. The 2-oxazoline ring is stable to mild reducing agents. Leffler and Adams (53) were able to reduce a number of 2-nitrophenyl-2-oxazolines to the corresponding 2-aminophenyl compounds with iron and hydrochloric acid at 100°C.

4. Oxidation

The 2-oxazolines are fairly stable toward oxidizing agents, as is illustrated by reactions of an oxidation type carried out with variously substituted 2-oxazolines. The ring is stable in these reactions toward nitrating mixtures or alkaline permanganate. Adams and Leffler (53) found it possible to nitrate 2-phenyl-2-oxazolines without cleaving the oxazoline ring. The ordinary nitric acid–sulfuric acid mixture was employed at a temperature not above 10°C. The effect of higher temperatures was not indicated. Billman and Parker (10) successfully oxidized 2-oxazolines containing a 4-methylol group to oxazolinecarboxylic acid according to the equation:



When R was ethyl or methyl, the oxazoline acid was isolated only with difficulty; isolation was easy when R was COOH.

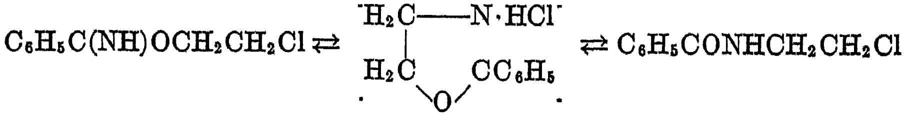
It is surprising to find no mention of conversion of a 2-oxazoline to an oxazole. By analogy with other "aromatic" compounds an instability of the dihydro compound with reference to the unsaturated compound is to be predicted. The failure to observe this transformation indicates a lower order of aromaticity in the oxazoles than other evidence indicates (89).

5. The iminoester–oxazoline–amide equilibrium

It has been previously noted that in their studies of the synthesis of oxazolines from iminoesters of β-chloroalcohols, Gabriel and Neumann had observed that the iminoester could be made to yield either the oxazoline or a β-chloroalkylamide. These authors regarded the transformation,

$$C_6H_5C(NH)OCH_2CH_2Cl \xrightarrow{\text{heat}} C_6H_5CONHCH_2CH_2Cl$$

as a simple rearrangement, but Wislicenus and Körber (90) present data indicating that an equilibrium exists, with the oxazoline as an intermediate:



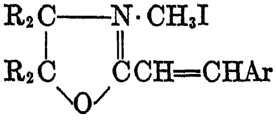
Wislicenus was led to this theory by a reaction of iminoesters previously studied by Wheeler and Johnson (87), which was essentially the same:

$$C_6H_5C(NH)OC_2H_5 + C_2H_5I \longrightarrow C_6H_5C(OC_2H_5)(NH\cdot C_2H_5I) \xrightarrow{100°C.} C_6H_5CONHCH_2CH_3 + C_2H_5I$$

Wislicenus regards Gabriel's reaction as a Wheeler type in which the use of the alkyl halide is rendered unnecessary by the presence of a CH_2CH_2Cl group in the molecule. The oxazoline corresponds to the intermediate in Wheeler's reaction. The validity of Wislicenus' mechanism is supported by several facts: (*1*) When an ether solution of the iminoester is concentrated, the residue contains the free oxazoline and the hydrochloride of the iminoester. The iminoester is more basic than the oxazoline and takes up the molecule of hydrogen chloride evolved during the reaction. (*2*) A possible equilibrium is indicated by the fact that when the iminoester hydrochloride and the free oxazoline are heated together, all of the ester does not go to the amide. (*3*) The oxazoline hydrochloride can be converted to the amide in 90 per cent yield by strong heat. The net result in the rearrangement of the iminoester to the amide is the migration of a haloalkyl group from oxygen to nitrogen; the similarity to the migration of an acyl group from oxygen to nitrogen discussed in connection with the hydrolysis of 2-oxazolines should be noted. It should also be pointed out that no attempt to convert an oxazoline hydrochloride back to the iminoester has yet been recorded.

6. Activity of a 2-methyl group

Since 2-methyl-2-oxazoline contains the grouping —$C(CH_3)=N$—, activation of the hydrogen atoms might be expected from analogy with α-picoline. Hamer and Rathbone (41) have found that the methiodide and ethiodide of oxazolines undergo typical condensations with aromatic aldehydes. Condensation occurs at the methyl group to yield compounds such as XXIX.



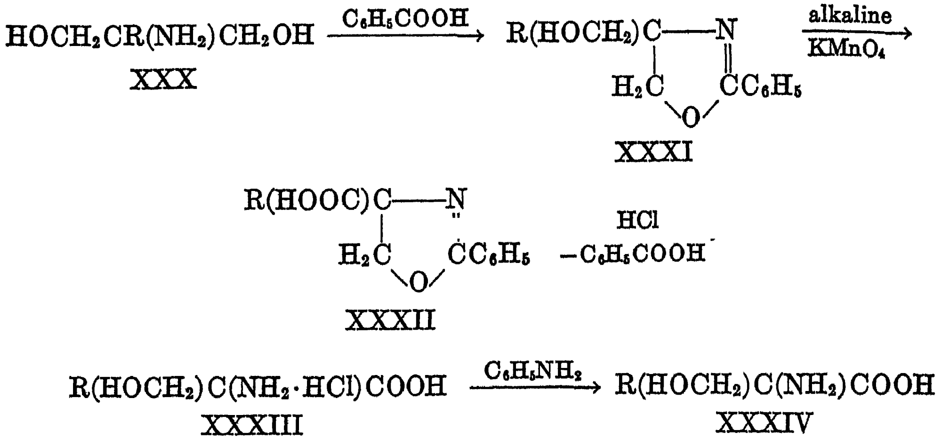
XXIX

Condensations of this type have not been studied by other investigators or with compounds other than the alkiodides of 2-methyl-2-oxazolines. Whether or not the quaternary salts lend an enhanced activity to the side-chain hydrogens is not known.

7. Synthetic applications

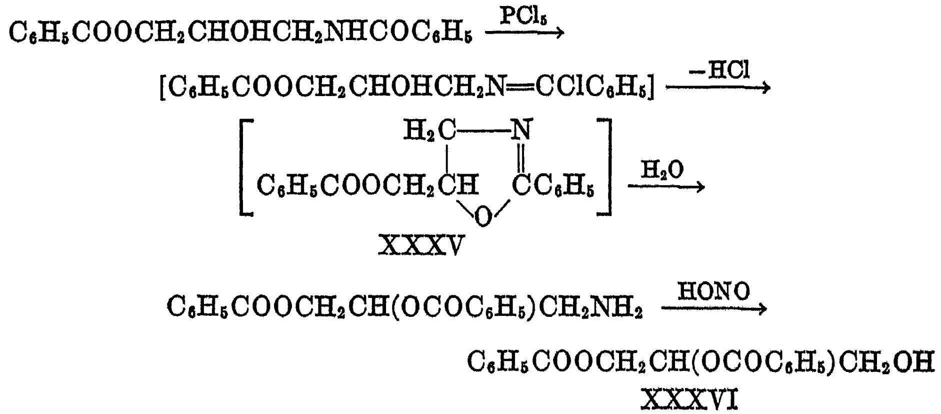
Oxazolines have been used as starting materials or intermediates in some interesting syntheses. Billman and Parker (10) have developed a process in which oxazolines are intermediates in the synthesis of α-amino-β-hydroxy acids. In searching for a synthetic route to serine, these investigators took advantage of the stability of the oxazoline ring toward mild oxidizing agents, and were thus able to oxidize one hydroxymethyl group in the oxazoline (XXXI) to a carboxyl group.

When R is CH_3, XXXI is formed in 67–69 per cent yield and XXXIV in 64.5 per cent yield from XXXI. Attempts to prepare serine by decarboxylation of the oxazoline (XXXI) or the amino acid (XXXIV) when R is COOH were unsuccessful.



XXX XXXI XXXII XXXIII XXXIV

It is possible that 2-oxazolines are formed as intermediates in a method proposed by Bergmann, Brand, and Dreyer (5) for preparation of diglycerides (XXXVI) whose acyl groups are in the 1- and 2-positions. The reactions are summarized as follows:


C₆H₅COOCH₂CHOHCH₂NHCOC₆H₅
PCl₅
[C₆H₅COOCH₂CHOHCH₂N=CClC₆H₅]
−HCl
H₂C—N
C₆H₅COOCH₂CH CC₆H₅
O
H₂O
C₆H₅COOCH₂CH(OCOC₆H₅)CH₂NH₂
HONO
C₆H₅COOCH₂CH(OCOC₆H₅)CH₂OH

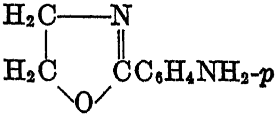
XXXV XXXVI

The oxazoline XXXV was not isolated, but since the net result of the reaction is an acyl migration from nitrogen to oxygen, it may be assumed as an intermediate. This work has been extended by Bergmann (3).

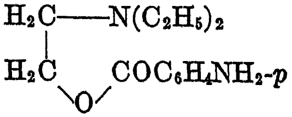
D. COMMERCIAL AND INDUSTRIAL USES OF 2-OXAZOLINES

Several 2-oxazolines have been considered for use as therapeutic agents. Because of their structural relationship to procaine (XXXVIII), 2-(aminophenyl)-2-oxazolines (XXXVII) might be expected to have some action as local anesthetics. Adams and Leffler, who prepared a series of these compounds (1, 53) and studied their action, found that they did induce local anesthesia. 2-(*m*-Aminophenyl)-2-oxazoline has about the same efficiency as procaine and is only about one-third as toxic. The utility of these compounds as anesthetics is limited by their low water solubility, which at the best is of the order of 1 per cent.

Quaternary salts of XXXVII, which might have increased water solubility, were not reported.

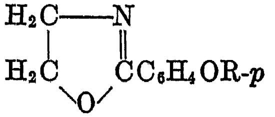


XXXVII

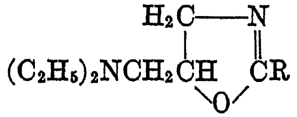


XXXVIII

In addition to the work of Adams and Leffler, Engelmann (26) has found oxazolines containing a *p*-alkoxyphenyl group in the 2-positions (XXXIX) to have action as local anesthetics, and Bockemühl and Knoll (14) have patented compounds such as XL as having therapeutic properties.



XXXIX



XL

$R = C_{17}H_{35}, C_{15}H_{31}$

A number of patents describe the use of oxazolines as wetting, emulsifying, or dispersing agents. Compounds used for these purposes always have a long aliphatic group in the 2-position, and are used in the form of salts (47); it has been found that salts of organic acids usually retain their surface activity longer than salts of inorganic acids. Wampner (84) has investigated compounds having hydroxymethyl groups in the 4-position. Compounds of this type have been used as cation-active finishing agents (63).

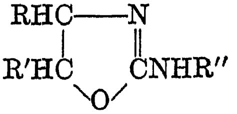
Lactic acid salts of oxazolines with long-chain groups in the 2-position have been used in cosmetic preparations (83). Photographic sensitizers of the cyanine type (41) and methine dyes (43) have been prepared with oxazoline nuclei. Ester amides, prepared from 4-mono- and 4,4-di-substituted-2-oxazolines and acid anhydrides by Tryon (78a), are stated to be useful as solvents or plasticizers and in the preparation of surface-active agents.

III. Substituted 2-Oxazolines

Of the many conceivable substituted 2-oxazolines, only a few are at all well known. These are the 2-amino, 2-mercapto, and a few hydroxy types. Some which have been listed in the tables in the previous section other than hydrocarbon substitution types are those containing carbethoxy, hydroxymethyl, carbethoxymethyl, and diethylaminomethyl groups. 2-Oxazolines containing vinyl, nitro, halo, aldehydo, or acetyl groups on the ring are apparently unknown. It was noted in Section I that substituted oxazolines, in which the substituent carries a hydrogen and is on a doubly bound ring carbon, may be in tautomeric equilibrium with oxazoles or oxazolidines. This is true of the 2-amino-, 2-mercapto-, and 2-hydroxy-2-oxazolines (X, XI, XII), which are the most widely studied of the 2-oxazolines. Similarities between the syntheses of these substituted types and unsubstituted 2-oxazolines were noted in Section IIA.

A. PSEUDOUREAS (2-AMINO-2-OXAZOLINES)

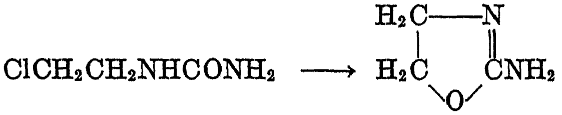
The pseudoureas, or 2-amino-2-oxazolines (XLI), are the most widely studied


XLI

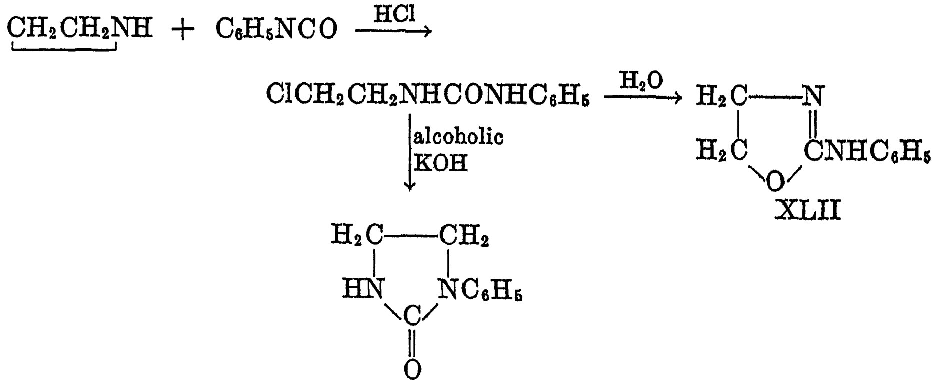
of the substituted 2-oxazolines. Equilibrium with the 2-iminoöxazolidine is possible, and the chemical properties indicate many reactions in both forms. The term "pseudourea" arises from the fact that these compounds are isomers of alkenylureas; thus 2-amino-2-oxazoline is isomeric with *N*-vinylurea and hence is "ethylene pseudourea"; 2-amino-5-methyl-2-oxazoline is isomeric with allylurea and is known as "propylene pseudourea". The compounds are solids and strong bases, forming well-characterized salts (58). 2-Phenylamino-2-oxazolines have been suggested for use as local anesthetics (64), but other possible applications have not developed from rather extensive theoretical studies.

1. *Syntheses of pseudoureas*

(*a*) *From β-haloalkylureas:* The hydrochloride of a pseudourea is obtained upon heating the haloalkylurea with water; the addition of alkali releases the free base.

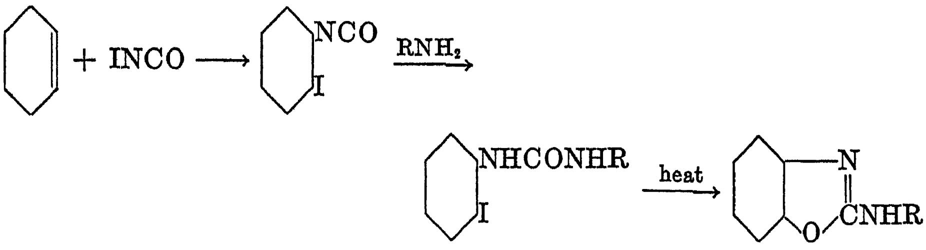


The yields vary from moderate to quantitative. The necessary unsymmetrically substituted ureas are available via several routes. One is the method of Takeda (75), who prepared pseudoureas by heating styrene dibromide, or substituted styrene dibromides, with urea. Gabriel (38) prepared unsaturated ureas from unsaturated amines and isocyanates; the addition of a halogen acid to the double bond then gave the requisite β-haloalkylurea. The product was cyclized to the pseudourea (XLII) on heating with water and to the imidazolidone (XLIII) with alcoholic potassium hydroxide.

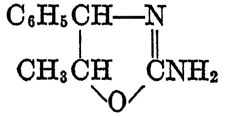

XLII

XLIII

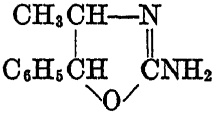
The haloalkylureas can be prepared by the addition of iodine isocyanate to an olefin to form a β-iodoisocyanate, which, when allowed to react with ammonia or an amine, yields the substituted urea. This method was studied by Birckenbach and Linhard (11, 12) in the synthesis of a series of cyclohexane derivatives.



These investigators also used olefins other than cyclohexene; however, the mode of addition of iodine isocyanate to unsymmetrical olefins is not known, and the position of groups in the pseudourea is then difficult to ascertain. Thus, they were able to convert *s*-phenylmethylethylene to the pseudourea in 94 per cent yield, but whether it had the structure XLIV or XLV was not determined; the compound was partially resolved, the *d*-form being isolated.



XLIV



XLV

Another route to the β-haloalkylureas is that originally used by Gabriel (32), in which a β-haloamine is allowed to react with cyanic acid. In this instance, the pseudourea may be obtained directly, without isolation of the intermediate β-halourea.

4-Keto derivatives of pseudoureas are formed by cyclization of α-haloacyl ureas, RCHBrCONHCONHR′, on treatment with alkali. Examples of such reactions have been reported by Aspelund (2a). He has also reported the conversion of 1,5-diphenyl-5-bromobarbituric acid to a pseudourea in a reaction which apparently involves formation of an α-haloacylurea as intermediate. Erlenmeyer and Kleiber (26a) report the formation of 5,5-diethyl-2-imino-4-oxazolidone from guanidine and ethyl α-hydroxy-α-ethylbutyrate.

(*b*) *From β-hydroxyalkylureas and thioureas:* The preparation of a pseudourea from a β-hydroxyalkylurea or thiourea requires a loss of water or hydrogen sulfide in the cyclization and is otherwise similar to the preparation from β-haloalkylureas. In using this method Söderbaum (72) cyclized a β-hydroxyalkylurea by heating with hydrochloric acid or by heating the thio analog with alcoholic mercuric oxide. The compounds prepared were derived from *s*-diphenylethanolamine through conversion to the urea or thiourea with isocyanates or isothiocyanates. The reaction was carried out where R, in XLVa, was H, CH_3, C_2H_5, C_6H_5, and *o*-$CH_3C_6H_4$.

TABLE 4

Pseudoureas: 2-amino-2-oxazolines

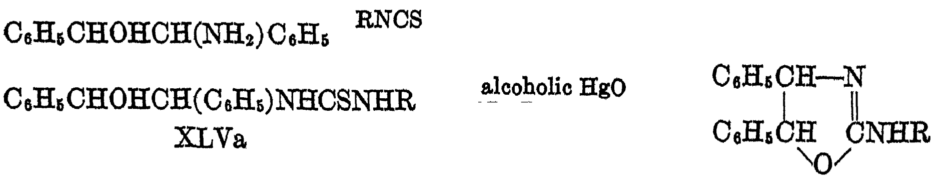
SUBSTITUENTS IN 2-AMINO-2-OXAZOLINE*			METHOD OF PREPARATION†	YIELD	REFERENCE
R	R′	R″			
				per cent	
			A;C	57	(11, 12, 32, 28)
	CH_3		C		(29)
	$HOCH_2$		A;C	33	(65, 28)
	ICH_2		A;C		(30, 65)
	$BrCH_2$		A;C		(30, 65)
	$ClCH_2$		C	33	(28)
	C_6H_5		A		(75)
$C_6H_5(CH_3)$	$CH_3(C_6H_5)$		A	94	(12)
p-$CH_3OC_6H_4$	CH_3		A		(75)
Cyclohexano			A	Quantitative	(12)
		C_6H_5	A		(38, 58)
	CH_3	p-$CH_3C_6H_4$	A		(58)
	CH_3	m-$CH_3C_6H_4$	A		(58)
	CH_3	$2,6$-$(CH_3)_2C_6H_3$	A		(58)
	CH_3	o-$CH_3OC_6H_4$	A		(58)
	CH_3	m-$C_2H_5OC_6H_4$	A		(58)
	CH_3	C_6H_5	A		(58)
	CH_3	p-$CH_3C_6H_4$	A		(58)
	$BrCH_2$	$BrCH_2CHBrCH_2$	A		(68)
	CH_3	CH_3, C_6H_5‡	A		(58)
Cyclohexano		C_6H_5	A	98	(12)
Cyclohexano		C_6H_5NH	A	89	(12)
C_6H_5	C_6H_5	R§	B		(72)

* R, R′, R″ of formula XLI; dash indicates R, R′, or R″ is H.

† A, from β-haloalkylurea; B, from β-hydroxyalkylurea or thiourea; C, from sodium cyanamide.

‡ Second hydrogen on nitrogen replaced by methyl.

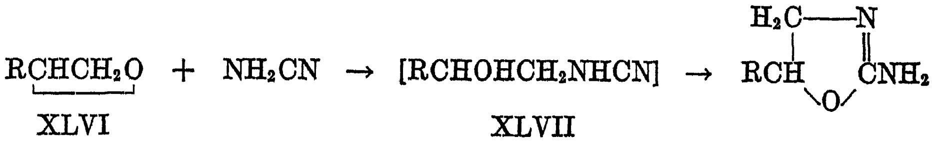
§ R is H, CH_3, C_2H_5, C_6H_5, o-$CH_3C_6H_4$.



XLVa

The reaction apparently takes a different course when —COOR replaces —CSNHR in XLVa. Thus, ethyl *N*-β-hydroxyethylcarbamate, $HOCH_2CH_2$-$NHCOOC_2H_5$, loses alcohol on heating to form an oxazolidone, not a 2-ethoxy-2-oxazoline (27a).

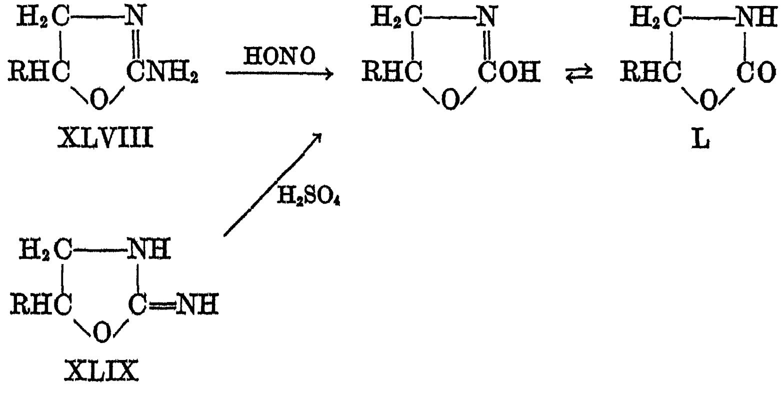
(*c*) *From sodium cyanamide and chlorohydrins:* Synthesis of pseudoureas (see table 4) from sodium cyanamide and chlorohydrins has no similarity to the syntheses of oxazolines as do the two preceding syntheses. Fromm and coworkers (28, 29, 30, 31) have described the reaction. The mechanism is not clear, but Fromm proposes that the sodium cyanamide reacts with the water present to form sodium hydroxide, which in turn dehydrohalogenates the chlorohydrin to the epoxide. Reaction of the epoxide (XLVI) with cyanamide then follows to form an unstable intermediate (XLVII), which rearranges to the pseudourea.



This mechanism is supported by the isolation of the pseudourea from the reaction when an epoxide is substituted for the chlorohydrin. Reaction between the chlorohydrin and sodium cyanamide to give the intermediate XLVII is not excluded by this evidence. The method has been successfully used with ethylene chlorohydrin and glycerol dichlorohydrin. A yield of 33 per cent of 2-amino-5-chloromethyl-2-oxazoline was reported from the epichlorohydrin. Reaction with chloroacetic acid gave an undescribed product.

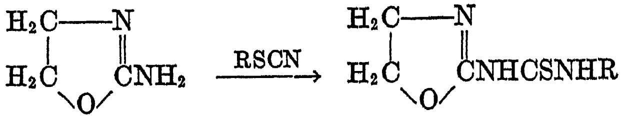
2. Reactions of pseudoureas

The tautomerism of pseudoureas has been extensively investigated, chiefly by Fromm (28, 29, 30), who has established that these compounds react either as 2-amino-2-oxazolines (XLVIII) or 2-iminoöxazolidines (XLIX).

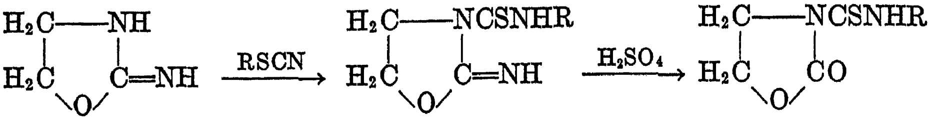


Thus pseudoureas are converted by heating with sulfuric acid to the oxazolidone (L); this can best be explained as hydrolysis of the imide form. This reaction was carried out with R = H, CH_2Cl, CH_3, and $CH_2SC_7H_7$. On the other hand, nitrous acid also produces the oxazolidone, and probably reacts with the amino form (XLVIII). This latter reaction has been carried out only with R = H, but there is no reason to doubt that it is general.

An independent line of investigation confirmed the existence of these tautomeric forms. Fromm (29, 30) found that isothiocyanates react with pseudoureas to form two thioureas, one high-melting and one low-melting. The low-melting compound (LIV) was assigned the imino structure, since sulfuric acid converts it to the oxazolidone (LV). The other compound (LIII) was then assumed to be a derivative of the primary amine (LI). The low-melting compound can be irreversibly converted to the high by heating.



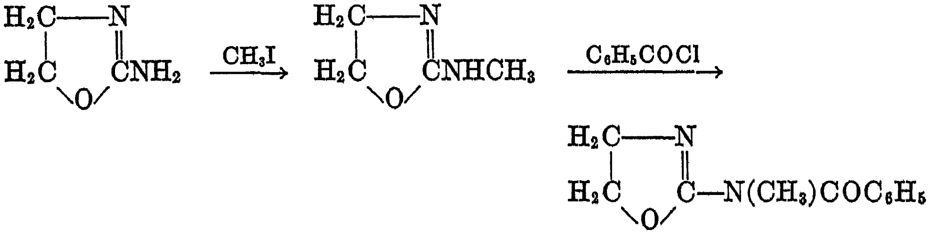
LI LIII



LII LIV LV

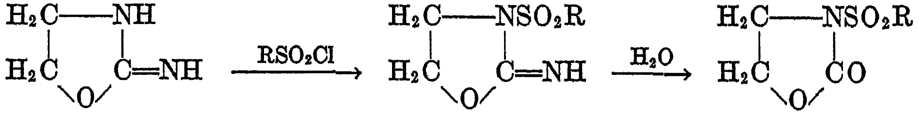
In connection with this study, Fromm (29) noted that 2-amino-4-phenyloxazole formed only one thiourea, indicating that in the oxazole series this tautomerism either does not exist or is very limited.

Other reactions of pseudoureas show that either the imino or the amino form can react (30). An *N*-methylpseudourea is obtained on reaction with methyl iodide. The benzoyl derivative (LVII) of the *N*-methylpseudourea is stable toward dilute acid and is therefore thought to be an *N*,*N*-derivative of the amino form and not a derivative of the imino form.



LVII

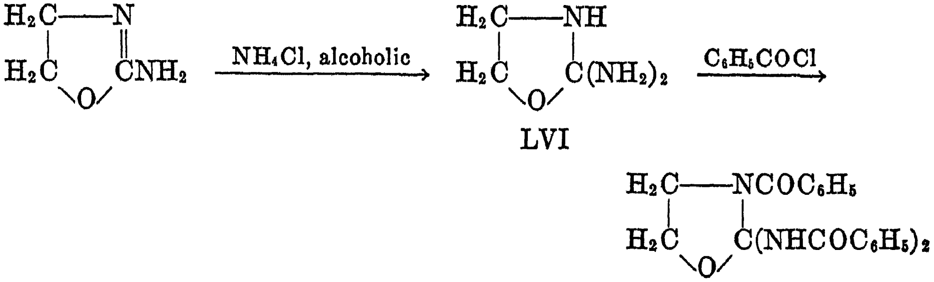
Acetylation with acetic anhydride and sodium acetate gives an acetyl derivative of the 2-amino-2-oxazoline form. The 2-alkylamino-2-oxazolines are similarly acetylated. Sulfonyl chlorides apparently attack either the ring nitrogen or both ring and side-chain nitrogen. In either case, the imino form is the one which reacts. Thus, benzenesulfonyl chloride forms the derivative LX, whose constitution was assigned (29) on the basis of the fact that it may be hydrolyzed to an oxazolidone (LXI).



LX LXI

Jensen (48) discovered that when acetylsulfanilyl chloride is used and the reaction carried out in pyridine, both nitrogen atoms are substituted; in acetone, however, only the 3-acetylsulfanilyl derivative was formed. Birckenbach and Linhard (12) report that hypobromite reacts with a pseudourea in the amino form to give an *N*,*N*-dibromo derivative.

A peculiar reaction with ammonia or amines has been described (28, 29). When a pseudourea is heated with ammonium chloride, or when its hydrochloride is heated with ammonia, addition to the double bond of the amino form occurs and a 2,2-diaminoöxazolidine (LVI) is formed (28, 29). Amines react similarly. The product forms a tribenzoyl derivative (LVII).


LVI

LVII

Rupture of the ring takes place under some conditions. Thus, at 100°C. with alcoholic ammonium chloride a guanidine derivative (LIX) is formed from 2-amino-5-chloromethyl-2-oxazoline (29).

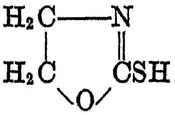
$$HCl\cdot H_2NC(=NH)NHCH_2CHOHCH_2Cl$$

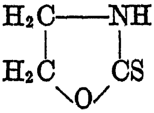
LIX

Rundqvist (65) has observed that sodium amalgam converts a 5-bromomethyl-2-amino-2-oxazoline to allylurea.

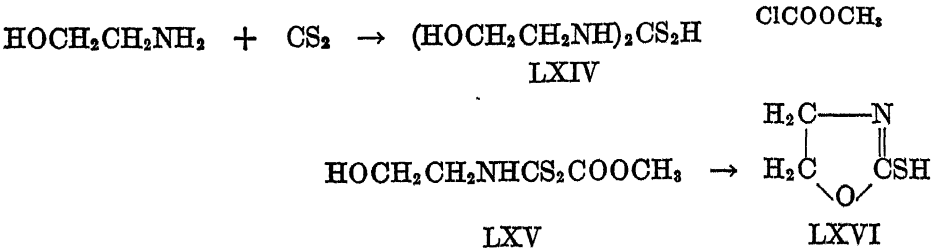
B. 2-MERCAPTO-2-OXAZOLINES

As the 2-amino-2-oxazolines, the 2-mercapto derivatives are capable of existing in the tautomeric oxazolidone form and react in either the thiol form (LXII) or the thione form (LXIII).


LXII


LXIII

2-Mercapto-2-oxazolines are formed by the action of carbon disulfide on β-hydroxyamines. Thus, Bruson and Eastes (20) converted 1-amino-2-hydroxy-2-methylpropane to 5,5-dimethyl-2-mercapto-2-oxazoline in almost quantitative yield. A similar method was employed by Sergeev (69), who found that β-hydroxyethylamine reacted first to form an intermediate (LXIV) which was converted by methyl chlorocarbonate to LXV, which upon long standing or upon heating formed the oxazoline (LXVI).



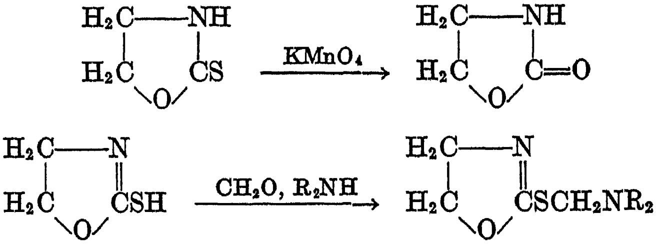
LXIV

LXV

LXVI

The formation of some 2-mercaptoöxazoline in the reaction between chlorohydrin and potassium thiocyanate has been reported (70).

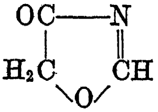
Reactions of both the thione form (LXII) and the thiol form (LXIII) are known. Permanganate converts the compound, presumably the thione form, to an oxazolidone (69). Iodine reacts, but the disulfide which might be expected was not isolated (60). Formaldehyde and amines replace the sulfhydryl hydrogen with a —CH_2NR_2 group (57).



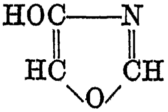
Various 2-mercapto-2-oxazolines have been evaluated for different uses. 2-Mercapto-2-oxazoline has antithyroid activity (60), and 5,5-dialkyl-2-mercapto-2-oxazolines have found some use in sprays and dusting powders for plants (19). A series of compounds were prepared by Mathes (57) through the action of formaldehyde and ammonia or amines on the 2-mercaptoöxazoline and studied as vulcanization accelerators. Of these, bis(5,5-dimethyl-2-oxazolin-2-ylmercaptomethyl)isopropylamine was said to have excellent properties as a vulcanization accelerator for rubber.

C. 4-KETO-2-OXAZOLINES

The 4-keto-2-oxazolines (LXVII) are tautomeric with 4-hydroxyoxazoles (LXVIII).

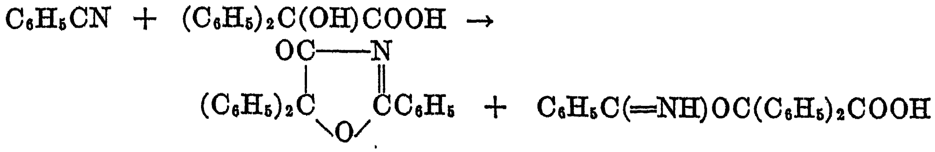


LXVII



LXVIII

The 4-ketopseudoureas (2a) are mentioned above. One other 4-keto-2-oxazoline has been described. The 2,5,5-triphenyl derivative (LXIX) was prepared by Japp and Findlay (45, 46) from benzonitrile and benzilic acid.

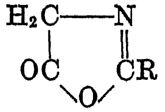


LXIX LXX

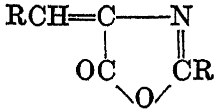
Benzimidoxydiphenylacetic acid (LXX) was also isolated from the reaction mixture and converted to the oxazoline by heating with acetic anhydride. The oxazolone is decomposed by heating with alkali or concentrated sulfuric acid. Ring closure of haloimides of the type $BrCH_2CONHCOR$ provides a source of 4-keto-2-oxazolines.

D. 5-KETO-2-OXAZOLINES

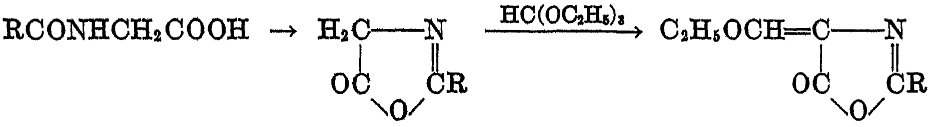
The 5-keto-2-oxazolines (LXXI) are the familiar azlactones. These are the α-acylamino acid inner anhydrides. Both the unsubstituted and 4-arylene types (LXXII) have been reviewed elsewhere (21) and need not be discussed here. Recent work (81a) on the synthesis of penicillin has started with various 4-hydroxymethylene-5-keto-2-oxazolines (LXXIIa) and their derivatives. The preparation of such derivatives from benzoyl- and some acyl-glycines, acetic anhydride, and ethyl orthoformate has been reported (2b). Others have been studied (24a, 81a). Those reported in the literature are listed in table 5. This preparation may take place through condensation of the active methylene group of the azlactone with ethyl orthoformate.



LXXI

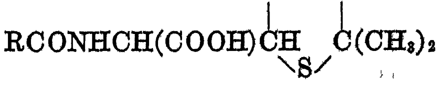


LXXII



LXXIIa

The hydroxymethylene compound itself is the enol form of the aldehyde. The objective of many attempts to synthesize penicillin (24a) has been to condense this aldehyde or one of its derivatives with penicillamine (β,β-dimethylcysteine). The hypothesis that the synthesis would proceed through thioacetal formation followed by ring closure to the thiazolidine is supported by known similar reactions. Thus, acetone reacts with cysteine to form 2,2-dimethyl-4-carboxythiazolidine (61a). The many unsuccessful attempts to achieve this synthesis and the extremely low yields finally obtained suggest the improbability of the reaction taking this course and of the oxazolone structure for penicillin. The unsuccessful anhydrization of the penicilloates is similar evidence indicating the



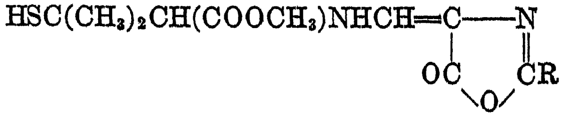
improbability of the oxazolone form (24a). A derivative of penicillin which has an oxazolone ring is formed on reaction of the methyl ester of penicillin with mercuric chloride (26b, 57a). This product, the methyl ester of penicillenic acid, has been assigned the structure:

TABLE 5

*4-Substituted-5-keto-2-oxazolines**

4-SUBSTITUENT	2-SUBSTITUENT	REFERENCE
$(CH_3)_2C=$	C_6H_5	(2b)
$CH_3OCH=$	C_5H_{11}	(81a)
$CH_3OCH=$	$C_6H_5CH_2$	(81a)
$CH_3OCH=$	$C_6H_5CH_2$	(24a, 81a)
$C_2H_5OCH=$	CH_3	(81a)
$C_2H_5OCH=$	C_5H_{11}	(2b)
$C_2H_5OCH=$	C_6H_5	(2b, 24a, 81a)
$C_2H_5OCH=$	$C_6H_5CH=CH$	(2b)
$C_2H_5OCH=$	p-$O_2NC_6H_4CH_2$	(81a)
$C_2H_5OCH=$	p-$O_2NC_6H_4CH=CH$	(81a)
$C_3H_7OCH=$	C_6H_5	(2b)
$HOCH=$	C_3H_7	(81a)
$HOCH=$	C_5H_{11}	(81a)
$HOCH=$	C_6H_5	(2b)
$HOCH=$	$C_6H_5CH_2$	(57a)
$HOCH=$	$C_6H_5CH=CH$	(2b)
$CH_3SCH=$	C_6H_5	(81a)
p-$CH_3CONHC_6H_4OCH=$	C_6H_5	(81a)

* *Added in proof:* See also *The Chemistry of Penicillin,* p. 743 ff. Princeton University Press, Princeton, New Jersey (1949).



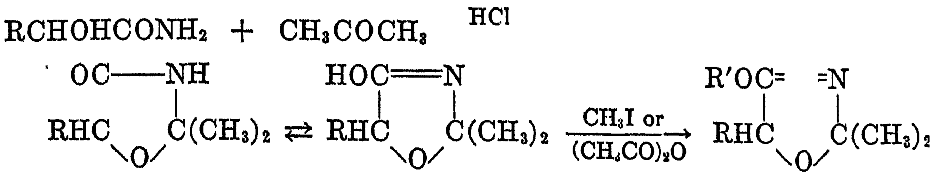
Hydrolysis of this ester with sodium hydroxide gives 4-hydroxymethylene-2-benzyloxazolone (57a).

IV. 3-OXAZOLINES

Neither 3-oxazoline itself nor any of its hydrocarbon derivatives are known. A report by H. O. L. Fischer, Dangschat, and Stettiner (27) contains the only known description of compounds with a 3-oxazoline ring.[3] These were obtained as *O*-methyl and *O*-acetyl derivatives (LXXIVa) of 4-oxazolidones (LXXIII). Zeisel alkoxyl determination established the *O*-methyl structure. Enolization to form a 3- instead of a 4-oxazoline ring was based on the similarity to the formation of *O*-alkyl derivatives of amides.

---

[3] *Note added in proof:* Others are reported in *The Chemistry of Penicillin,* Princeton University Press, 1949, pp. 739–42, as pseudoxazolones.



LXXIII LXXIV LXXIVa

The 4-oxazolidones were prepared from amides of mandelic, glycolic, and lactic acids. This synthesis is similar to the synthesis of oxazolidines from aldehydes or ketones and β-hydroxyamines (58a). The wide applicability of the oxazolidine synthesis suggests that this synthesis of 4-oxazolidones and their derivative 4-hydroxyoxazolines can be extended.

TABLE 6

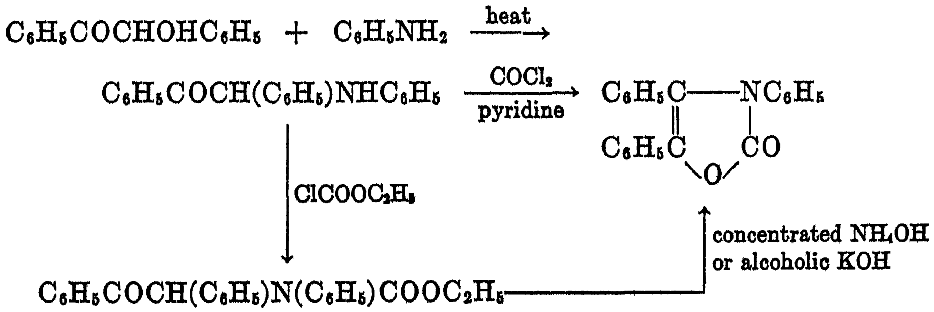
*4-Oxazolines**

3-SUBSTITUENT	4-SUBSTITUENT	5-SUBSTITUENT	YIELD
			per cent
C_6H_5	C_6H_5	C_6H_5	80
o-$CH_3C_6H_4$	C_6H_5	C_6H_5	
m-$CH_3C_6H_4$	C_6H_5	C_6H_5	
p-$CH_3C_6H_4$	C_6H_5	C_6H_6	
C_6H_5	H	C_6H_5	Almost quantitative
o-$CH_3C_6H_4$	H	C_6H_5	60
p-$CH_3C_6H_4$	H	C_6H_5	
m-$CH_3C_6H_4$	H	C_6H_5	60
$C_{10}H_7$	H	C_6H_5	

* Reported by McCombie and coworkers (55, 56).

V. 4-OXAZOLINES

4-Oxazoline is not known and neither are any of its simple hydrocarbon derivatives. The only known types are the 3,5-diaryl and 3,4,5-triaryl-2-keto-4-oxazolines prepared by McCombie and coworkers (55, 56) by the following series of reactions:



N-Phenacylaniline, $C_6H_5COCH_2NHC_6H_5$, gave 3,5-diphenyl-2-keto-4-oxazoline. Others are included in table 6.

These compounds are highly stable substances. 3,4,5-Triphenyl-2-keto-4-oxazoline is unchanged by several hours refluxing with aqueous acid, alcohol, or phosphorus pentachloride. As would be expected from the negativity of the *N*-substituent, this substance is not sufficiently basic to yield a hydrochloride or chloroplatinate. Typical ketone reactions are not shown; heating in a sealed tube with phenylhydrazine gives no phenylhydrazone. Reduction with sodium amalgam gives no clear-cut results; the oxazolone is sometimes reduced to bibenzyl and sometimes recovered unchanged. The hydrogen atom of carbon number four, in compounds having no 4-substituent, cannot be replaced by bromine.

VI. REFERENCES

(1) ADAMS, R., AND LEFFLER, M. T.: U. S. patent 2,114,326 (April 19, 1938); Chem. Abstracts **32**, 4726 (1938).
(2) ANDREASCH, R.: Monatsh. **5**, 33–46 (1884).
(2a) ASPELUND, H.: Acta Acad. Aboensis, Math. et Phys. **12**, No. 5, 33 pp. (1939) (Pub. 1940); Chem. Abstracts **41**, 2413 (1947); **33**, 6801 (1939); **35**, 2143 (1941).
(2b) BARBER, H. J., SLACK, R., STICKINGS, C. E., ELLIOTT, D. J., AND ATTENBURROW, J.: British patent 585,089 (January 30, 1947).
(3) BERGMANN, M.: Z. physiol. Chem. **137**, 27–46 (1924).
(4) BERGMANN, M., AND BRAND, E.: Ber. **56B**, 1280–3 (1923).
(5) BERGMANN, M., BRAND, E., AND DREYER, F.: Ber. **54B**, 936–65 (1921).
(6) BERGMANN, M., BRAND, E., AND WEINMANN, F.: Z. physiol. Chem. **131**, 1–17 (1923).
(6a) BERGMANN, M., AND GRAFE, K.: Z. physiol. Chem. **187**, 196 (1930).
(7) BERGMANN, M., AND MIEKELEY, A.: Z. physiol. Chem. **140**, 128–45 (1924).
(8) BERGMANN, M., AND SABETAY, S.: Z. physiol. Chem. **137**, 46–61 (1924).
(9) BETTZIECHE, F.: Z. physiol. Chem. **146**, 227–40 (1925).
(10) BILLMAN, J. H., AND PARKER, E. E.: J. Am. Chem. Soc. **67**, 1069–70 (1945).
(11) BIRCKENBACH, L., AND LINHARD, M.: Ber. **64B**, 961–8 (1931).
(12) BIRCKENBACH, L., AND LINHARD, M.: Ber. **64B**, 1076–87 (1931).
(13) BLACKBURN, S., MIDDLEBROOK, W. R., AND PHILLIPS, H.: Nature **150**, 57 (1942).
(14) BOCKEMÜHL, M., AND KNOLL, R.: U. S. patent 1,958,529 (May 15, 1934); Chem. Abstracts **28**, 4539 (1934).
(15) BOOKMAN, S.: Ber. **28**, 3111–21 (1895).
(16) BROMUND, W. H., AND HERBST, R. M.: J. Org. Chem. **10**, 267–76 (1945).
(17) BROOKER, L. G. S.: German patent 719,270 (March 5, 1942); U. S. patent 2,165,692 (July 11, 1939); Chem. Abstracts **37**, 1723 (1943); **33**, 8213 (1939).,
(18) BRÜHL, J. W.: Z. physik. Chem. **16**, 193–241 (at pp. 209, 218) (1895).
(19) BRUSON, H. A.: U. S. patent 2,282,931 (May 12,1942); Chem. Abstracts **36**, 6301(1942).
(20) BRUSON, H. A., AND EASTES, J. W.: J. Am. Chem. Soc. **59**, 2011–13 (1937).
(21) CARTER, H. E.: *Organic Reactions*, Vol. III, Chap. 5. John Wiley and Sons, Inc., New York (1946).
(22) DERSIN, H.: Ber. **54B**, 3158–62 (1921).
(23) DIELS, O., AND BECCARD, E.: Ber. **39**, 4125–32 (1906).
(24) DILTHEY, W., AND FRIEDRICHSEN, J.: J. prakt. Chem. **127**, 292–306 (1930).
(24a) EDITORIAL BOARD OF PENICILLIN MONOGRAPH: Science **105**, 653–9 (1947).
(25) ELFELDT, P.: Ber. **24**, 3218–28 (1891).
(26) ENGELMANN, M.: U. S. patent 2,027,031 (January 7, 1936); Chem. Abstracts **30**, 1519 (1936).
(26a) ERLENMEYER, H., AND KLEIBER, A.: Helv. Chim. Acta **21**, 111–13 (1938).
(26b) EVERS, N., AND CAMPBELL, N. R.: Ann. Repts. Soc. Chem. Ind. on Progress Applied Chem. **30**, 502–4 (1945).

(27) Fischer, H. O. L., Dangschat, G., and Stettiner, H.: Ber. **65B,** 1032–7 (1932).

(27a) Franchimont, A. P. N., and Lublin, A.: Rec. trav. chim. **21,** 45–55 (1902).

(28) Fromm, E., Barrenscheen, H., Frieder, J., Pirk, L., and Kapeller, R.: Ann. **442,** 130–49 (1925).

(29) Fromm, E., Kapeller-Adler, R., Friedenthal, W., Stangel, L., Edlitz, J., Braumann, E., and Nussbaum, J.: Ann. **467,** 240–74 (1928).

(30) Fromm, E., Kapeller, R., Pirk, L., Hahn, A., and Leipert, T.: Ann. **447,** 259–84 (1926).

(31) Fromm, E., and Honold, E.: Ber. **55,** 902–11 (1922).

(32) Gabriel, S.: Ber. **22,** 1139–54 (1889).

(33) Gabriel, S.: Ber. **22,** 2220–3 (1889).

(34) Gabriel, S., and Colman, J.: Ber. **47,** 1866–73 (1914).

(35) Gabriel, S., and Eschenbach, G.: Ber. **30,** 2494–7 (1897).

(36) Gabriel, S., and Heymann, T.: Ber. **23,** 2493–2502 (1890).

(37) Gabriel, S., and Neumann, A.: Ber. **25,** 2383–8 (1892).

(38) Gabriel, S., and Stelzner, R.: Ber. **28,** 2929–38 (1895).

(39) Gabriel, S., and Stelzner, R.: Ber. **29,** 2381–91 (1896).

(40) Granächer, C.: Helv. Chim. Acta **8,** 865–73 (1925).

(41) Hamer, F. M., and Rathbone, R. J.: British patent 541,330 (November 24, 1941); Chem. Abstracts **36,** 4767 (1942).

(42) Hopkins, C. Y.: Can. J. Research **16B,** 341–4 (1938).

(43) I. G. Farbenindustrie A.-G.: French patent 848,028 (October 24, 1939); Chem. Abstracts **35,** 5716 (1941).

(44) Jänecke, E.: Ber. **32,** 1095–1103 (1899).

(45) Japp, F. R., and Findlay, A.: Proc. Chem. Soc. **15,** 165–6 (1899).

(46) Japp, F. R., and Findlay, A.: J. Chem. Soc. **75,** 1027–32 (1899).

(47) Jayne, D. W., Jr., and Day, H. M.: U. S. patent 2,329,619 (September 14, 1944); Chem. Abstracts **38,** 1304 (1944).

(48) Jensen, K. A.: Dansk. Tid. Farm. **16,** 1–10 (1942); Chem. Abstracts **37,** 4375 (1943).

(49) Jones, G. D.: J. Org. Chem. **9,** 484–99 (1944).

(50) Kay, P.: Ber. **26,** 2848–52 (1893).

(51) Krabbe, W., Eisenlohr, W., and Schöne, H.: Ber. **73B,** 656–60 (1940).

(52) Krabbe, W., Polzin, E., and Culemeyer, K.: Ber. **73B,** 652–5 (1940).

(53) Leffler, M. T., and Adams, R.: J. Am. Chem. Soc. **59,** 2252–8 (1937).

(54) Loder, D. J.: U. S. patent 2,402,198 (June 18, 1946); Chem. Abstracts **40,** 5457 (1946).

(55) McCombie, H., and Parkes, J. W.: J. Chem. Soc. **101,** 1991–8 (1912).

(56) McCombie, H., and Scarborough, H. A.: J. Chem. Soc. **103,** 56–63 (1913).

(57) Mathes, R. A.: U. S. patent 2,343,566 (March 7, 1944); Chem. Abstracts **38,** 3163 (1944).

(57a) Committee on Medical Research, O.S.R.D., Washington, and The Medical Research Council, London: Science **102,** 627–9 (1945); Nature **156,** 766–7 (1945).

(58) Menne, E.: Ber. **33,** 657–65 (1900).

(58a) Meyer, V., and Jacobson, P.: *Lehrbuch der organischen Chemie*, Vol. 2, Part 3, p. 521. Walter de Gruyter and Co., Berlin and Leipzig (1923).

(59) Michels, W.: Ber. **25,** 3048–55 (1892).

(60) Miller, W. H., Roblin, R. O., Jr., and Astwood, E. B.: J. Am. Chem. Soc. **67,** 2201–4 (1945).

(61) Nagai, W. N., and Kanao, S.: Ann. **470,** 157–82 (1929).

(61a) Ratner, S., and Clarke, H. T.: J. Am. Chem. Soc. **59,** 200–6 (1937).

(62) Rehländer, P.: Ber. **27,** 2154–61 (1894).

(63) Robinette, H.: Am. Dyestuff Reptr. **31,** No. 24; Proc. Am. Assoc. Textile Chemists Colorists, p. 575–80 (1942); Chem. Abstracts **37,** 1608 (1943).

(64) Rose, C. L., Schonle, H. A., and Chen, K. K.: Pharm. Arch. **11,** 81–9 (1940); Chem. Abstracts **35,** 1522 (1941).

(65) Rundqvist, C.: Arch. Pharm. **236,** 445–7 (1898); J. Chem. Soc. **76,** 16

(66) SALOMON, A.: Ber. **26**, 1321–30 (1893).
(67) SAULMANN, F.: Ber. **33**, 2634–9 (1900).
(68) SCHMIDT, E.: Arch. Pharm. **236**, 441–5 (1898); J. Chem. Soc. **76**, 16 (1899).
(69) SERGEEV, P. G., AND IVANOVA, S. N.: J. Gen. Chem. (U.S.S.R.) **7**, 1495–1500 (1937); Chem. Abstracts **32**, 2534 (1938).
(70) SERGEEV, P. G., KOLYCHEV, B. S., AND KONRAT'EV, A. G.: J. Gen. Chem. (U.S.S.R.) **7**, 2600–4 (1937); Chem. Abstracts **32**, 2534 (1938).
(71) SOCIÉTÉ POUR L'INDUSTRIE CHIMIQUE À BÂLE: French patent 830,125 (July 21, 1938); Chem. Abstracts **33**, 1416 (1939).
(72) SÖDERBAUM, H. G.: Ber. **28**, 1897–1903 (1895).
(73) STRAUSS, E.: Ber. **33**, 2825–30 (1900).
(74) TAKEDA, J.: J. Pharm. Soc. Japan No. 426, 691–709 (1917); Chem. Abstracts **11**, 3241 (1917).
(75) TAKEDA, J., AND KURODA, S.: J. Pharm. Soc. Japan **449**, 561–608 (1919); Chem. Abstracts **14**, 179 (1920).
(76) TIETZMAN, J. E., DOHERTY, D. G., AND BERGMANN, M.: J. Biol. Chem. **151**, 387–94 (1943); Chem. Abstracts **38**, 1210 (1944).
(77) TRYON, P. F.: U. S. patent 2,372,409 (March 27, 1945); Chem. Abstracts **39**, 4894 (1945).
(78) TRYON, P. F.: U. S. patent 2,372,410 (March 27, 1945); Chem. Abstracts **39**, 4336 (1945).
(78a) TRYON, P. F.: U. S. patent 2,410,318 (October 29, 1946); Chem. Abstracts **41**, 2076 (1947).
(79) UEDINCK, A.: Ber. **32**, 967–79 (1899).
(79a) VALCO, E. I.: U. S. patent 2,416,552 (February 25, 1947); Chem. Abstracts **41**, 3823 (1947).
(80) VANGHELOVICI, M., AND STEFANESCU, A.: Soc. Chim. România Sect. Soc. române Stiinte, Bul. Chim. pură apl. [2] **3A**, 159–81 (1941–42); Chem. Abstracts **38**, 5501 (1944).
(81) VANGHELOVICI, M., AND MOISE, I.: Soc. Chim. România Sec. Soc. române Stiinte, Bul. chim. pură apl. [2] **3A**, 85–106 (1941–42); Chem. Abstracts **38**, 5500 (1944).
(81a) DU VIGNEAUD, V., CARPENTER, F. H., HOLLEY, R. W., LIVERMORE, A. H., AND RACHELE, J. R.: Science **104**, 431–3 (1946).
(82) VARGA, E., AND FODOR, G. v.: J. prakt. Chem. **150**, 94–8 (1938).
(83) WAMPNER, H. L.: U. S. patent 2,368,078 (January 23, 1945); Chem. Abstracts **39**, 5416 (1945).
(84) WAMPNER, H. L.: U. S. patent 2,402,791 (June 25, 1946); Chem. Abstracts **40**, 5273 (1946).
(85) WENKER, H.: J. Am. Chem. Soc. **57**, 1079–80 (1935).
(86) WENKER, H.: J. Am. Chem. Soc. **60**, 2152–3 (1938).
(87) WHEELER, H. L., AND JOHNSON, T. B.: Am. Chem. J. **21**, 185–93 (1899); **23**, 135–50 (1900).
(88) WHITE, T.: J. Chem. Soc. **1940**, 428–37.
(89) WILEY, R. H.: Chem. Revs. **37**, 401–42 (1945).
(90) WISLICENUS, W., AND KÖRBER, H.: Ber. **35**, 164–8 (1902).
(91) WOLFHEIM, F.: Ber. **47**, 1440–52 (1914).
(92) ZEMPLEN, G., GERECS, A., AND RADOS, M.: Ber. **69**, 748–54 (1936).
(93) ZEMPLEN, G., GERECS, A., AND ILLES, E.: Ber. **71**, 590–6 (1938).

SOLUBILITY MEASUREMENT IN THE CRITICAL REGION

HAROLD SIMMONS BOOTH AND RICHARD MACPHERSON BIDWELL[1]

Morley Chemical Laboratory, Western Reserve University, Cleveland, Ohio

Received April 6, 1948

CONTENTS

I. Historical summary 478
II. Nature of the critical state 480
  A. Pure substances 480
  B. Critical curves 480
  C. Influence of solubility 481
  D. Limited solubility 482
  E. Multicomponent systems 484
  F. Compound formation 484
III. Geological significance 484
  A. Nature of magma 485
  B. Orthomagmatic stage 485
  C. Pneumatolytic stage 485
  D. Hydrothermal stage 486
  E. Influence of volatile constituents 486
IV. Engineering significance 488
V. Experimental methods 490
  A. Introduction 490
  B. Hydrothermal synthesis 490
    1. Refinements of hydrothermal methods 490
  C. Direct solubility measurements 491
    1. Sampling method 491
    2. Internal filters 491
    3. Sampling of solutions near the critical point 492
    4. Synthetic method 493
  D. Physical measurements 493
    1. Spectrometric measurements 493
    2. Conductivity measurements 493
  E. Loss of weight method and vapor-phase transport 494
  F. Balancing method 494
  G. Agitation 495
  H. Materials 495
  I. Miscellaneous investigations 496
VI. Systems reported in the literature 497
  A. Introduction 497
  B. Individual systems 497
    1. Systems with ammonia 497
    2. System: carbon dioxide–iodine 497
    3. Systems: carbon dioxide with various solutes 498
    4. Systems with carbon disulfide 498
    5. System: ethane–hexachloroethane 499
    6. Systems: ethanol–inorganic compounds 499
    7. Systems: ethanol–organic compounds 501

---

[1] Holder of a Harshaw Chemical Company Fellowship for Fundamental Research in Chemistry, 1946–49.

8. System: ethyl ether–anthraquinone 501
9. Systems with ethyl ether 501
10. Systems: hydrogen–alkali metals 502
11. Systems: methanol–inorganic salts 502
12. Systems with sulfur dioxide 502
13. Systems: water–inorganic compounds 502
14. System: water–silica 505
C. Electrical properties of solutions 507
VII. Conclusion 509
VIII. References 510

I. HISTORICAL SUMMARY

The earliest observations of critical phenomena were made in 1822 by Cagniard de la Tour (17), who first observed the disappearance of the vapor–liquid interface, characteristic of the critical state. In 1869 Andrews (4), in his classical work on carbon dioxide, established the continuity of the gaseous and liquid phases. The thermodynamic theory of the critical phase for one- and two-component systems was carefully worked out by van der Waals (112, 113) in 1881 and later elaborated by others (88, 95, 96, 97).

Studies demonstrating the possibility of continuous change of phase from liquid to gas led Hannay and Hogarth (48, 49, 50, 51, 52, 53) to inquire whether solvent action is characteristic only of the liquid phase, or whether gases might not also be capable of dissolving solids. Their work, including the study of solutions of several inorganic salts in ethanol, demonstrated that the solution of salts is not limited to the liquid phase. In particular, it was noted that an alcoholic solution of cobalt chloride above its critical point had essentially the same absorption spectrum as in the normal liquid state, thus indicating that the usual ionic condition persists even above the critical temperature of a solution. The concentrations of solid in the supercritical phase were far higher than could be accounted for by the normal volatility of the salts concerned. Whether this is a true solution or an enhanced volatility is perhaps more a matter of nomenclature than of theory (75, 106, 118).

The observations of Hannay and Hogarth, though fairly conclusive and verified by others (19, 119, 120), were widely misinterpreted, particularly by Ramsay (86), with the result that they were largely ignored. The extreme experimental difficulties associated with such complicated critical studies discouraged extensive investigation of this field, and not until the work of Centnerszwer (20, 21, 22, 23, 124) between 1903 and 1910 was a systematic quantitative study made of the critical points of solutions of salts. Earlier observations had, however, been made of the behavior of mixtures of gases near the critical point (11, 18).

The experiments of Centnerszwer demonstrated that the presence of dissolved salts normally causes a rise in the critical point of a liquid, analogous to the rise in the normal boiling point and similarly proportional to salt concentration. The presence of dissolved material may either raise or lower the critical point of a liquid, depending on the nature of the solution. The critical point may thus, like the melting point, be used as a criterion of chemical purity (20).

Other determinations from this viewpoint are those of Büchner (16) and of Schröer (92), who reached similar conclusions, though differing somewhat in their quantitative interpretation. Schröer observed wide variations in the behavior of different salts in aqueous solution when the critical region was reached.

In 1912 Niggli (78) published the first of his works pointing out the importance of critical phenomena in geological processes, particularly as regards the formation of ores and rare minerals. This work was later expanded into a complete system of the phase relations likely to occur in natural magmas, and application of these relationships to the explanation of the formation of actual minerals (77). Such work, combined with the experimental synthesis of minerals and the study of the phase relations of magmatic components, carried on particularly at the Geophysical Laboratory of the Carnegie Institution of Washington, has added greatly to understanding of the processes resulting in the formation of ores and other characteristic end-products of magmatic activity. A large proportion of this work is not particularly concerned with critical phenomena, and hence is not within the scope of this paper. The interpretation of many of the results in terms of critical occurrences (77, 96, 97) has, however, been invaluable in the clear understanding of observations.

It seems wise to limit this paper to the type of system where, of two components, the triple point of one lies at a considerably higher temperature than the critical point of the other. This is characteristic of solutions of many salts in water or in organic liquids, as well as of the complex mixtures which make up magmas, the several components of the latter being readily classifiable as volatile and refractory substances, respectively. On the more extreme types, exemplified by combination of a "permanent gas" with a metal, metallic oxide, or other such refractory material, little conclusive information seems to be available. Systems of more nearly comparable critical constants have been reviewed elsewhere (3). These include the industrially very important mixtures of petroleum hydrocarbons, as well as the systems carbon dioxide–permanent gases (11, 64) and argon–boron trifluoride (13).

This review presents a general survey of the field from a chemical point of view. Earlier reviews of related nature (73, 75) are more concerned with hydrothermal and pneumatolytic studies of minerals.

It has been demonstrated that the solubility of salts in alcohols is not discontinuous at the critical point of the solutions. It is also shown that solubility in the single-phase region is strikingly dependent upon pressure, or perhaps more directly the density or "concentration" of the solvent (figures 4, 6, 7, 8, 9) (106, 118).

Conductivity studies (30, 33, 45, 46, 47, 63, 64, 124) substantiate solubility measurements on salt–alcohol systems and in addition show that the dielectric constants of solutions may not suffer large changes at their critical points, but may still permit dissociation of ionic salts to give electrically conducting gas-phase solutions (figures 10, 11). The recorded curves of conductivity as a function of pressure and temperature show a close resemblance to the solubility variation under similar conditions (111). This is related to the fact that the processes of solution and ionization depend to some extent upon the same solvent characteristics.

II. NATURE OF THE CRITICAL STATE

A. PURE SUBSTANCES

For any sufficiently stable substance there exists a temperature, with associated pressure and specific volume, at which the liquid and vapor phases become identical. This point, known as the critical point, is often spoken of as the temperature above which a substance can exist only as a gas. The material existing in a state not too remote from the critical state may, however, have nearly as many of the characteristics, from the conventional viewpoint, of the liquid as of the gaseous state. As was shown by Andrews (4) and later corroborated and developed from a theoretical point of view (112), a transition may be made from attenuated gas to cold liquid, passing through this region, without at any point encountering any discontinuous change of phase, as for example the formation of mist droplets, bubbles, or a meniscus. Thus water vapor or superheated steam may be heated to 400°C., above the critical temperature, compressed isothermally under several hundred atmospheres pressure, and then cooled to cold water, without condensation.

Although a substance above the critical point may be predominantly of a gaseous nature, it may still retain many of the attributes of the liquid phase. It may have comparable density, and may be capable of dissolving solids to about the same extent as liquid of that density, and of forming ionic solutions. It is perhaps not strictly correct to speak of solution properties as typical of liquids, in view of the data indicating solution and conductivity in the vapor phase even below the critical point (30, 33, 45, 46, 47, 63, 106, 111). "Solubility above the critical point," as used in this paper, refers always to solubility above the critical point of the *solution*, that is, in the gas phase.

That such dispersion is true solution in the gas phase, and not volatilization, has repeatedly been stated, although the distinction is not entirely clear-cut. If volatilization, it is at least associated with an abnormally high vapor pressure. The quantity of solute found in the gas phase is frequently far greater than could be accounted for by normal volatility (103, 106, 118). The effect of pressure on the amount of solid dispersed is much larger than that expected from normal increase of vapor pressure due to external pressure.

The spectra of such solutions have been observed to be the same as for solutions of the same solute in the usual liquid phase, and to differ from those obtained on vaporization of pure solid substances (19, 50, 51, 119). Ionic solutions of salts conduct electric current above their critical temperatures (30, 33, 45, 46, 47, 63, 64).

B. CRITICAL CURVES

It was mentioned earlier that the presence of dissolved salts causes a rise in the critical temperature of a liquid. This rise is actually (75) only the beginning of a curve connecting the critical points of solvent and solute, and representing the critical point at each intermediate concentration (curve A_cB_c, figure 1). Points A_c and B_c are the respective critical points of the pure substances. The other lines of the diagram give pressure–temperature relations under conditions of equilibrium between different phases characteristic of this system of a volatile component A and a refractory component B. "B + L + V", for example, refers to equilibrium between pure solid B, a liquid phase, and a vapor phase; that is, a solution saturated with respect to B, under its own vapor pressure. Such solubility curves run to the eutectic *E* from the respective triple points A_t and B_t of the two components. Diagrams similar to figure 1 may be drawn up, relating concentration to pressure, temperature, and specific volume (73, 77).

The critical points for mixtures represent, in general, neither the maximum temperature nor the maximum pressure attainable in the presence of liquid when a mixture of two components in the proper proportions is heated, but are the true critical points, i.e., the point for each mixture where the liquid and vapor phases become identical. For pure substances, the points of maximum temperature and maximum pressure coincide with the critical point, but for mixtures these points are in general all different, permitting the occurrence of retrograde phenomena (3). For solutions of salts in ordinary solvents, the critical curves show a maximum pressure at some intermediate composition and temperature, the geological importance of which is to be pointed out. This is, however, not invariably the case.

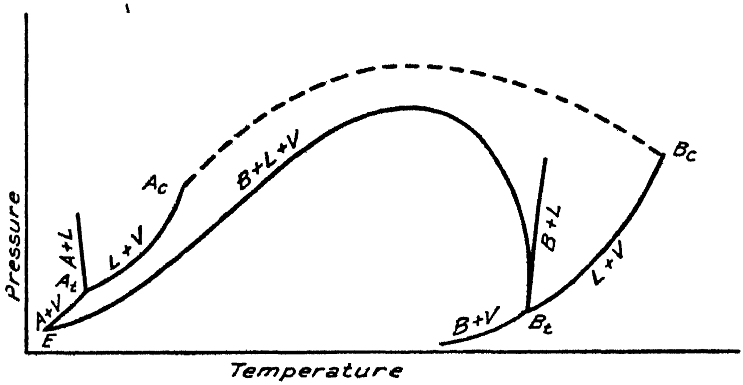


FIG. 1. Pressure–temperature diagram for a two-component system without critical end-points (after Morey and Ingerson (73)).

C. INFLUENCE OF SOLUBILITY

In the case of very soluble substances, critical phenomena may never occur for saturated solutions, that is, in the presence of solid solute. This is a result of the raising of the critical temperature by the dissolved substance, which becomes increasingly soluble as the temperature rises, allowing further increases in the temperature of liquid solution, until the composition of pure solute is attained. This is the type of system, exemplified by water–silver nitrate, of which the phenomenon of "second boiling point" is characteristic; the total vapor pressure of the mixture falling again, in general, as the composition of pure salt is approached (66). In this case the *solubility curve*, representing equilibrium between solid solute and solution, lies entirely below the critical curve, as in figure 1. The limitation of critical phenomena to unsaturated solutions under these conditions is obvious.

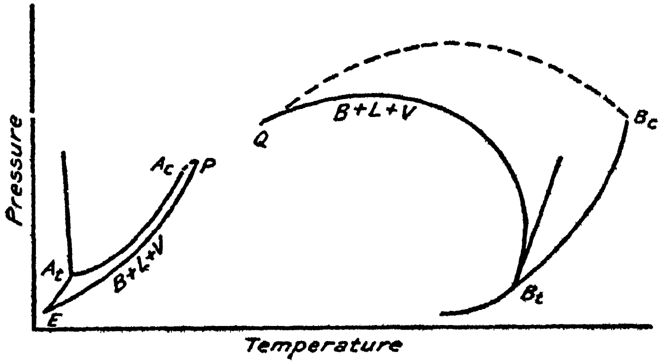


FIG. 2. Pressure–temperature diagram for a two-component system having two critical end-points (after Morey and Ingerson (73)).

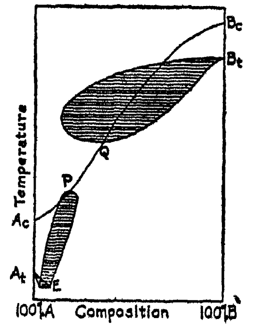


FIG. 3. Temperature–composition diagram for a two-component system having two critical end-points (after Morey and Ingerson (73)).

D. LIMITED SOLUBILITY

Of more interest in this work is the case where solubility is more limited, critical phenomena arise, and solubility in the supercritical "fluid" phase enters into consideration. This frequently occurs when the solubility of the solute falls with rising temperature. In this case the solubility curve cuts the critical curve twice, and in the intervening region both curves are discontinuous, as shown in figure 2. Two critical solution points P and Q exist, representing the upper and lower temperature limits, respectively, of two regions where equilibrium between liquid and vapor phases exists, these equilibria being represented in figure 3 by the isothermal tie-lines connecting the two phases in equilibrium with solid at the given temperatures in the loops EP and B_tQ, respectively. In the intervening region, solid component B is in equilibrium with a single fluid phase, having certain characteristics of both liquid and gas. In this phase the solubility of component B depends upon both temperature and pressure, as may best be shown (figure 4) by the experimental curves of Smits (96) for the system ether–anthraquinone: an example, under comparatively mild conditions, of the type of behavior expected of sparingly soluble solids in solution in water and other common solvents. The two loops appear, as in the general temperature–concentration diagram (figure 3), representing equilibria between liquid and vapor phases in the presence of solid anthraquinone at various temperatures. n the intermediate region of a single fluid phase isobars have been drawn in, howing the solubility of anthraquinone in ether in this phase, as dependent pon pressure. It should be obvious, however, that a rigid designation of comonents as solvent and solute, respectively, may lead to confusion; outside the

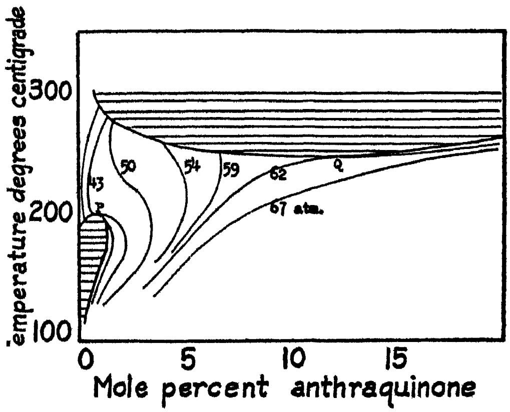


FIG. 4. Temperature–composition diagram for the ether–anthraquinone system, showing isobars (after Smits and Niggli (77)).

ower loop of these diagrams the choice of "solvent" is often dubious. The actual situation is that of reciprocal solubilities; it would be least ambiguous to speak only of relative proportions of the two components in each phase.

Other similar systems have been explored to a lesser extent than the ether–anthraquinone system, and some preliminary study (103, 118) has shown parallel behavior in a small region immediately above the critical point for the system silica–water.

In each of the two loops characteristic of these temperature–composition diagrams, three-phase equilibrium is represented between vapor and liquid phase solutions and solid B, three phases in all, with compositions fixed at the temperature in question, in accordance with the phase rule. Between P and Q, where only a single fluid phase exists in contact with solid, divariance (variation in composition with both temperature and pressure) is to be expected. In the small region A_1E (figure 3), solid component A is in equilibrium with solution and vapor, this being the region of the familiar lowering of the freezing point of a solvent.

These temperature–composition diagrams do not, of course, represent isobaric sections, but show conditions at pressures allowing multiphase equilibria. A pressure variation similar to that shown in figure 5 is characteristic of the three-phase portions of systems such as those shown in figures 3 and 4.

E. MULTICOMPONENT SYSTEMS

Where more than two components are present in a system, the relations involved are somewhat more complex, owing to the greater number of variables. No phenomena occur, however, which are intrinsically different from those already described, and those considered in elementary phase study.

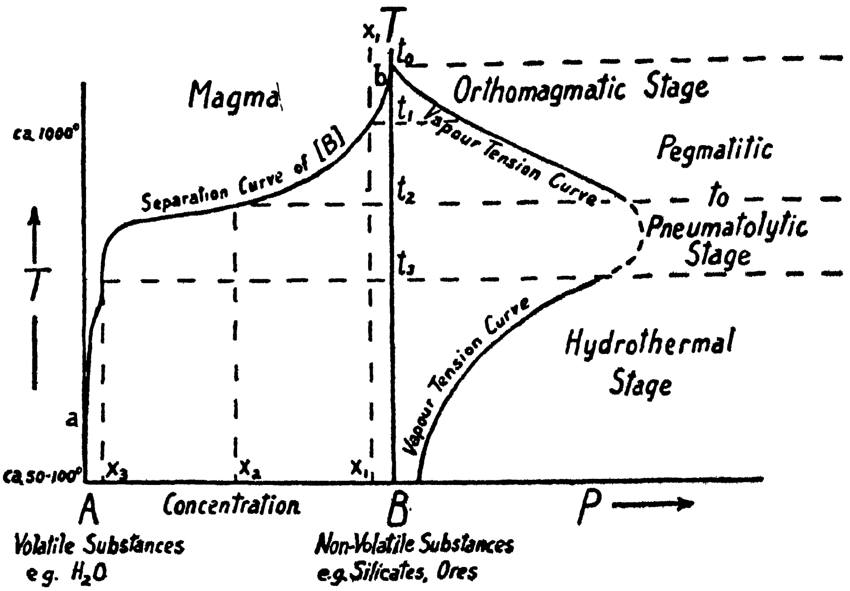


FIG. 5. Temperature–concentration and temperature–pressure diagrams for a cooling ideal magma (Niggli (77a)).

F. COMPOUND FORMATION

Under certain conditions, two or more components present in a system may interact to form compounds, as for example the various silicates and hydrates formed in the system SiO_2–K_2O–H_2O (69). An interesting example of compound formation, as occurring in systems of the type described in this paper, is the CaO–CO_2 system, which, though well known, has not generally been considered from this viewpoint.

III. GEOLOGICAL SIGNIFICANCE

A large proportion of the information on phase relations between volatile and refractory substances is the result of studies of the formation of minerals and ores, that is, from a geological viewpoint. This is related to the fact that only under the tremendous external pressure which can be exerted by hundreds of feet of rock can appreciable proportions of volatile matter be kept in intimate contact with refractory substances near their melting points, except under artificial conditions. Only the development of modern metals of high tensile strength has permitted the study of such systems.

A. NATURE OF MAGMA

A magma may be considered as a mass of fluid material in the earth's crust consisting principally of the common rock-building elements, with traces of practically all other elements in some form or other. Originally it can be considered as being an essentially homogeneous mass at something over 1000°C. Appreciable proportions of volatile materials, particularly water, hydrogen chloride, hydrogen fluoride, carbon dioxide, hydrogen sulfide, silicon tetrafluoride, and certain of their reaction products with less volatile components, are present in the normal magma, being held in solution under pressure.

B. ORTHOMAGMATIC STAGE

As the magma loses heat to the surrounding rock, the highest melting major components, in general, begin to crystallize out. As a result, the concentration of volatile components in the remaining solution rises. The total vapor pressure of the system also rises, the increase due to the increased proportion of volatile materials being greater than the decrease resulting from the reduction of temperature; there is thus observed an increasing pressure as the temperature falls (see figures 3 and 5). This is merely the generalized aspect of "second boiling point." Under these conditions, a phase containing a large proportion of volatiles tends to separate in the magma, the system now entering the state described by the upper loop of figure 3. The principal processes occurring here are crystallization of such rocks as granites, and simultaneous separation of "contact gas." This gas may (*a*) escape through fissures or other openings in the surrounding rock, (*b*) remain included in cavities in the newly formed rock, (*c*) be occluded in solid solution, (*d*) react chemically, or (*e*) redissolve in the remaining liquid magma (121). This stage in magmatic evolution, known as the orthomagmatic stage, is characterized by the presence of molten rock-forming material, and is not entirely different from early concepts of the interior of the earth.

One important result of the increase in pressure as the magma cools is the expansion of the containing space, resulting often in cracks in the surrounding rock, thus permitting escape of the more mobile phases, which may eventually produce such formations as veins, beds, and dikes. If it occurs close enough to the surface, volcanic eruption may result.

C. PNEUMATOLYTIC STAGE

As the magma cools further, the proportion of volatile material in both fluid phases becomes quite high, in agreement with figure 3; then eventually a portion of the siliceous material redissolves in the "contact gas," and the two phases become identical at the point corresponding to the upper region critical point Q (figures 2 and 3). This and the region of a single fluid phase, immediately below, constitute the region of *pneumatolytic* processes, where a considerable amount of chemical activity and transport of dissolved material is facilitated by the extremely high pressure and the action, chemically and as a vehicle, of the low-viscosity contact gas or supercritical fluid phase. The actual mechanism of the separation of phases has been a subject of considerable dispute (14).

In the cooling of the magma, those substances crystallize out first with respect to which the solution first becomes saturated. Since a few rock-building minerals are present in overwhelming proportions (99 per cent of all igneous material is said to be included under only ten mineral types), these compounds crystallize out first, along with certain isomorphic substances, concentrating rare and valuable ores and minerals in the mother liquors. These rare minerals, being by virtue of their low concentration still in solution in the pneumatolytic stage, are often carried away from the main mass of the magma, the concentration continuing during transport, and are eventually deposited in veins, beds, or dikes. Pegmatite dikes are intrusive masses of coarse-grained granite formed during this stage, usually containing an appreciable proportion of rare minerals, which tend to concentrate further during deposition of the granite, giving associated, comparatively pure deposits of these minerals. Such minerals are consequently often found in pegmatites or in fissures in the surrounding rocks. The large size characteristic of pegmatite crystals is in harmony with the concept of their deposition from relatively dilute solution.

D. HYDROTHERMAL STAGE

On further cooling, the magma enters the final *hydrothermal* stage, in which liquid and gaseous phases are again present together, but now with a greater preponderance of volatile material. The processes occurring in this stage are solution, deposition, and reaction in an essentially aqueous medium, containing, however, significant proportions of other volatiles. Minerals are deposited from dilute solution under these conditions, large deposits being built up by continued flow of the cooling saturated solutions. Hydrothermal reactions, being closest to ordinary chemical experience, are comparatively well understood, and many of the postulated or proven reactions will be included in the next section.

E. INFLUENCE OF VOLATILE CONSTITUENTS

As suggested earlier, mineral deposition depends to a considerable degree on the action of volatile substances, of which perhaps the most important are water, carbon dioxide, hydrogen sulfide, hydrogen chloride, hydrogen fluoride, sulfur dioxide, and silicon tetrafluoride (25, 77, 127). These substances are known as "mineralizers," because of their action in promoting mineral concentration and crystallization during the solidification of rock-forming material, particularly in pegmatite dikes. They, or more commonly their ultimate reaction products, such as hydrated minerals, carbonates, sulfides, halides, and oxides, are found as constituents of, or associated with, minerals and ores, frequently those of commercial value. Thus, for example, fluorspar is often found with sphalerite (ZnS), galena (PbS), and chalcopyrite ($CuFeS_2$).

Volcanoes, fumaroles, and hot springs are all surface manifestations of magmatic activity, and represent the ultimate separation and dispersal of volatile constituents. The sampling of substances from these sources gives valuable information toward the better understanding of geological processes, although it must be understood that the ejected material is in many cases heavily contaminated with foreign substances (127). Gases such as those mentioned above, associated often with well-defined solid minerals, have frequently been detected in these exhalations (1, 93, 126, 127, 128).

Mineralizers may include, however, not only such highly volatile substances, but a complete range of compounds, ranging from the most volatile to such materials as fluorspar, with melting points above 1000°C. The characteristic behavior of mineralizers is a type of solvent action, associated in some cases with chemical reaction. Any substance can be a mineralizer which appreciably increases the mobility of a system, allowing more rapid attainment of equilibrium and thus promoting the separation of pure phases. Fluorides, in fact, owe their name to the viscosity-lowering properties of fluorspar, commonly used as a flux.

The extent of the influence of fluorides, characterized as "*mineralisateurs par excellence,*" on geological processes has perhaps been underestimated because of the difficulty of determining gaseous fluorides in samples, both because of their high reactivity, particularly on glass, and because of the difficulty of precise analysis in the presence of other reactive gases and of each other. Large proportions of fluoride gases were collected at the fumaroles of the Valley of Ten Thousand Smokes (126, 127, 129).

Many systems and reactions have been studied, and others postulated, for the formation of minerals, particularly through the action of mineralizers as reagents (128). A few examples are given below (1, 26, 77, 127).

$$\begin{gathered}
Fe_2Cl_6 + 3H_2O \rightleftharpoons 6HCl + Fe_2O_3 \\
Fe_2Cl_6 + 2H_2O \rightleftharpoons 4HCl + 2FeOCl \\
Fe_2Cl_6 \rightleftharpoons Fe_2Cl_4 + Cl_2 \\
6Fe_2O_3 \rightleftharpoons 4Fe_3O_4 + O_2 \\
4HCl + O_2 \rightleftharpoons 2H_2O + 2Cl_2 \\
SnCl_4 + 2H_2O \rightleftharpoons 4HCl + SnO_2 \\
TiF_4 + 2H_2O \rightleftharpoons 4HF + TiO_2 \\
CaF_2 + H_2O \rightleftharpoons CaO + 2HF \\
BeCl_2 + H_2O \rightleftharpoons BeO + 2HCl \\
2AlF_3 + 3H_2O \rightleftharpoons 6HF + Al_2O_3 \\
3SiF_4 + 2Al_2O_3 \rightleftharpoons 3SiO_2 + 4AlF_3 \\
SiO_2 + 4HF \rightleftharpoons SiF_4 + 2H_2O \\
PbCl_2 + H_2S \rightleftharpoons PbS + 2HCl \\
4AsCl_3 + 2CoCl_3 \rightleftharpoons Co_2As_4 + 9Cl_2 \\
FeSO_4 + H_2S + S \rightleftharpoons FeS_2 + H_2SO_4 \\
3FeO + H_2O \rightleftharpoons Fe_3O_4 + H_2 \\
3S + 2H_2O \rightleftharpoons 2H_2S + SO_2 \\
H_2S + 2H_2O \rightleftharpoons 3H_2 + SO_2 \\
2H_2S \rightleftharpoons 2H_2 + S_2 \\
S_2 + 4CO_2 \rightleftharpoons 2SO_2 + 4CO \\
2NaCl + H_2O \rightleftharpoons 2HCl + Na_2O
\end{gathered}$$

Many of these are noticeably high-temperature reactions, particularly if equilibrium is regarded as occurring in the presence of appreciable proportions of the several substances concerned. There may be some doubt whether magmatic emanations often contain low enough concentrations of water to permit the extensive transfer of such elements as aluminum and titanium in the form of their halides.

It has been stated that the gaseous emanations from certain incandescent lavas may consist principally of sodium chloride (28). How generally this may be true, however, is in doubt (1).

There is increasing acceptance of the concept of simple solution in volatile substances, frequently in the gas phase, as the principal medium in the transfer and concentration of ore minerals, rather than the formation of volatile compounds. These solutions are generally primarily aqueous, though the presence of other compounds acting as solvent is not excluded, but at lower temperatures water is usually present in far larger proportion than other potential solvents (56).

With respect to the occurrence of solutions above their critical points, it is believed that the normal concentration of mineralizers of moderate volatility (alkali and alkaline earth halides, for example) in normal magmatic mixtures may be high enough to make true critical phenomena indeed of rare occurrence in the presence of the main body of the magma, the liquid phase of the magma then consisting, as it cools, of solutions in a series of increasingly volatile mixtures of compounds, but with no principal solvent substance above its critical point, as altered by the presence of other components, at any time. This does not necessarily apply, however, to fluid "emanations" separated from the magma, which, with their dissolved solids, are probably frequently above the critical temperature.

The potentialities of sodium chloride as a solvent, or mineralizer, are shown by its use as such in industrially practicable syntheses (10, 11b, 11c, 11d). Fluorides have frequently been used in syntheses difficult or impossible in the absence of mineralizing action (27, 125).

Although much information has been acquired through the study of minerals and volcanic activity, no complete picture of mineral-forming processes can be built up until a great deal of further experimental study has been made of solutions and reaction equilibria at temperatures and pressures similar to those encountered in nature.

IV. ENGINEERING SIGNIFICANCE

It has frequently been observed that a solid deposit appears on steam turbine blades, in spite of precautions to prevent entrainment (24, 65, 70, 104, 105, 106). These deposits are observed to occur particularly at regions of falling pressure, such as the convex sides of turbine blades, and consist principally of quartz, amorphous silica, and sodium salts.

It has been the object of some study to prevent the formation of these deposits, since the efficiency of a turbine is appreciably impaired by their presence. The deposition of silica is particularly undesirable, since it requires the use of either caustic or mechanical cleaning for removal.

Reviews and criticisms have been written (70, 106) on the limited number of studies made, from an engineering point of view, of this problem. In some cases (24, 70, 106) turbines were examined, and the nature and quantity of the deposits ascertained. An effort was made to correlate these observations with operating conditions.

Of a somewhat more fundamental nature are experiments in which saturated or superheated steam is passed over dry solids or saturated solutions, and the effluent vapor analyzed (104, 106). In other experiments steam from a boiler containing known concentrations of salts was analyzed, the primary objective being the prevention of the carrying over of silica by adjustment of operating conditions and addition of salts to form compounds less easily dissolved in steam (106).

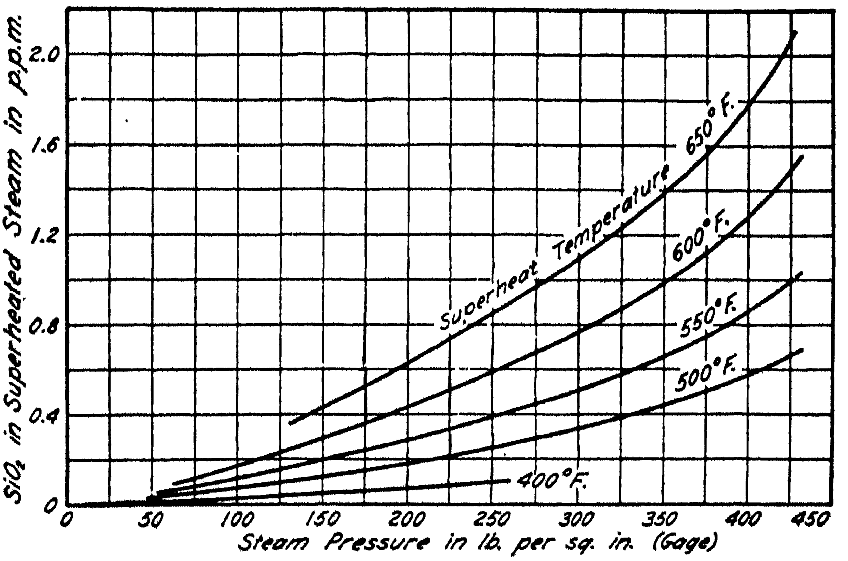


FIG. 6. Solubility of silica in superheated steam as a function of temperature and pressure (Straub (106)).

Since these experiments were carried out under such conditions as to exclude entrainment except as a known very minor source of error, it is established that many substances have a measurable solubility in steam (70, 106). This solubility, first becoming prominent in the region of 50 atm., rises sharply with increasing pressure (figure 6), agreeing in this respect with measurements by different techniques above the critical point in the silica–water (118), anthraquinone–ether (97), and other systems (23, 97, 111). The high solubility of silica, or silicic acid, in both liquid and gaseous water as the critical region is approached is noteworthy.

V. Experimental Methods

A. INTRODUCTION

Owing to the experimental difficulties encountered, studies involving critical conditions have been very limited in number and in scope. Studies of critical phenomena of mixtures as such have been largely limited to non-aqueous systems, because of the extreme corrosiveness of water at those temperatures, which is not likely to be decreased by the presence of dissolved salts. Ordinary construction metals and most glasses are attacked by water near its critical point.

B. HYDROTHERMAL SYNTHESIS

A large proportion of the work involving solutions above the critical point has been from the viewpoint of the hydrothermal synthesis and alteration of minerals. These procedures involve, in general, the heating of a mineral or a known mixture of pure chemicals with water and perhaps other reagents at the temperature of interest until equilibrium has been attained, the purpose being to simulate natural geological processes. Some type of pressure vessel is, of course, usually necessary as a container. The pressure vessel and contents are cooled rapidly, often quenched, to room temperature, preserving the phases formed under the experimental conditions. The crystals and glass (representing the liquid phase) from the bomb are analyzed chemically and microscopically. The techniques and history of these investigations, extending back to 1845, have been reviewed (55, 73, 75).

From the various combinations of phases found stable at various temperatures, pressures, and proportions of components, a complete phase diagram describing the system can be drawn up. Many of these reactions are carried out at temperatures well above the critical point of water. Over a hundred minerals have thus been synthesized.

As one possible variation, carbon dioxide may be introduced into the hydrothermal bomb as an additional component, the action of carbonic acid at lower temperatures simulating the weathering of rocks (81), as well as taking part in syntheses.

The results of hydrothermal reactions and more direct solubility measurements may supplement each other in the complete analysis of a system.

1. Refinements of hydrothermal methods

Much of the more recent hydrothermal work (74) has been done with a highly complex apparatus in which steam is introduced into the reaction vessel from an external boiler. The outer parts of the container are kept at a desired temperature, say 500°C., and an internal furnace heats the charge, wrapped in platinum foil, to the test temperature. The charge is suspended from fusible wires and is quenched by simply passing a current through the wires, which then allow it to fall into the cooler lower part of the container. It will be seen that great flexibility is attainable by the use of such an apparatus. Other gases, as well as steam, may of course be introduced at controlled temperatures and pressures.

For the study of the oxidation of minerals in water, a bleed-off valve attached to a hydrothermal bomb permitted the escape of evolved hydrogen, thus adding another variable to the simulation of magmatic conditions (31).

C. DIRECT SOLUBILITY MEASUREMENTS

1. Sampling methods

Since the direct chemical analysis of the contents of a container at high pressures presents unusual experimental difficulties, it is desirable, in order to determine the solubility of a substance under such conditions, to withdraw a portion of the saturated solution from the bomb. Since the solution is exposed to considerable changes in both temperature and pressure upon leaving the bomb, it is usually necessary to separate the solution from the undissolved solid phase *before* it leaves the container. This requires, in general, some type of filter *inside* the pressure vessel.

A popular technique is the intersection of a sintered filter stick through the top of a pressure vessel to a point near the bottom. Such apparatus was used to measure the solubility of calcium sulfate (107) in water to temperatures somewhat below the critical. The liquid in this case was drawn off into a smaller pressure vessel, originally at atmospheric pressure.

For measurements of the solubility of silica to 360°C., a filter in the base of the container, consisting of graded particles of the solute, was used. The solution was drawn off into a water-cooled condenser, a procedure said to be usable only where the solute is not deposited rapidly from supersaturated solution, or where it may be removed easily from surfaces (55).

2. Internal filters

In many cases it may be expected that an appreciable change in the constitution of the solution will occur on passing through the filter, resulting in deposition of solute in and near the filter itself. In work close to the critical point, appreciable changes may also occur in the nature of the solution remaining inside the pressure vessel, even while sampling. It is often of interest to examine this material in the equilibrium state.

To effect the separation of solution from solid phases under equilibrium conditions, the *internal filter* was developed (108). In this type of apparatus the container is divided essentially into two sections by a filter or screen, the solid solute being restricted to one side. During the time equilibrium is being attained, usually with agitation, the solution is on the lower side of the partition, in contact with the solid phase. Filtration is effected by simply inverting the container, the solution then passing through the filter. Subsequent cooling may cause part of the solute to precipitate, but the entire contents, solid and liquid, of that part of the bomb will in general give the solubility at the equilibrium temperature. With this type of apparatus, unlike some sampling devices, only one sample may be taken for each filling of the bomb.

A disadvantage of internal filters is the difficulty with which vapor passes through pores blocked with liquid or solid. This may perhaps often be alleviated by tilting the container slowly enough so that only part of the filter is, at any time, immersed in the solution. An improved design (60, 61, 62) has a small tube inserted through the filter, itself equipped with a filter disk. The displaced gas rises through this tube unhindered and passes out through the small filter, under an appreciable hydrostatic head, into the now uppermost compartment.

A greatly refined and modified design of internal filter (72) includes a stirrer which, when equilibrium has been attained, is lowered to cut a gold diaphragm. The charge above the diaphragm can then pass through to a filter crucible. Steam pressure in the upper chamber is externally controlled, and that below made approximately the same, by the presence of a small, calculated amount of water placed there at the beginning of the experiment.

3. Sampling of solutions near the critical point

Sampling methods and internal filters are applicable to solubility measurements, in general, which are made appreciably below the critical point. Above the critical, however, there exists only a single homogeneous fluid phase, and there is consequently no driving force to cause the liquid to pass through an internal filter, unless possibly some external force is applied to one side of the filter. In the case of external sampling, this is essentially the situation. The application of force, however, changes the state of the system, and equilibrium conditions then no longer exist. While this is of little importance at ordinary temperatures and pressures, a negligible amount of solvent being evaporated or condensed, with a corresponding small temperature change, in the critical region and above a large change in the density of solvent occurs under external influences. The solubility of a solid in this region has been shown (97, 106, 114, 118) to vary considerably with pressure, and therefore with the density of the solvent.

Under these conditions, it is to be expected that as soon as a portion of the solution is removed, the remaining solution will become appreciably supersaturated. It is then likely that a part of the dissolved salt will precipitate, either in the container or in the pores of the filter itself. The extent of this deposition may be difficult to estimate.

Where pressure is applied to an internal filter, the compressed fluid will in general tend to dissolve more solute. This will often lead to negligible error, owing to the slowness of the dissolving process.

Apparatus not necessarily subject to these difficulties has been developed (8, 122, 123). In these the sampling chamber is an integral part of the pressure vessel.

Both upper and lower phases of a two-phase mixture can be sampled by a technique involving displacement with mercury. In one such apparatus (55a) two smaller bombs were connected at the sides of the main pressure vessel. Tubes connected the tops of these two bombs with, respectively, the upper and lower parts of the large chamber. The lower ends of the two were both connected through tubes, closed by valves, to the bottom of the larger container. When a valve was opened, mercury flowed down into the main chamber, displacing a sample of its contents into the space formerly occupied by mercury, in the top of one of the small bombs. Solubilities above the critical point are, of course, thus measurable. Since there is no change in pressure or volume, the equilibrium conditions are maintained while sampling. Filters could, if necessary, be inserted in the lines.

4. Synthetic method

One of the more successful procedures for the determination of solubilities in closed systems, where transparent containers may be used, is the synthetic method (21, 111). Known quantities of solvent and solute are enclosed in the system, and, sufficient agitation being supplied, the points are observed at which the solute is completely dissolved and at which it begins to crystallize out again. These temperatures can be very precisely observed, in the case of an anisotropic solute, by the use of polarized light, the crystals shining brilliantly against a dark field (11a). The synthetic technique has been successfully used above the critical point (23, 95, 111).

D. PHYSICAL MEASUREMENTS

Although chemical determinations on solutions above the critical point are very difficult, certain physical methods have more promise. The methods which have been used in this and related fields are (*1*) spectrometric or colorimetric and (*2*) electrical conductivity.

1. Spectrometric measurements

The observation of visible color and qualitative spectroscopic examination of solutions of cobalt chloride, ferric chloride, a chlorophyll extract, and alizarin in ethanol (7, 50, 51, 52, 53), and of iodine in carbon dioxide (19, 119) were used to indicate the presence and physical state of dissolved material in liquid, vapor, and "fluid" phases. A similar technique has been used on aqueous solutions of potassium dichromate, copper sulfate, cobalt chloride, and cupric chloride (7, 92). These methods have apparently not yet been used for precise quantitative measurement.

2. Conductivity measurements

The use of electrical conductivity measurements for the determination of solubilities of sparingly soluble salts in ionizing solvents has been recommended (82, 87). Although excellent conductivity and dielectric constant measurements have been made above the critical point on solutions of known concentration (30, 33, 45, 46, 47, 63, 64, 124), the somewhat more tedious determination of solubility in this region by such means has apparently not been carried out.

For solutions in alcohols, ammonia, and sulfur dioxide at the critical point, glass capsules with sealed-in electrodes were found satisfactory. For other, more corrosive, solvents, such as water and hydrogen fluoride, metal apparatus modelled after that used by A. A. Noyes and his colleagues (82) for their classic studies on conductivity and ionization of water and aqueous solutions at elevated temperatures should be suitable.

The limitation of most chemical thought to "beaker chemistry" is well demonstrated by the fact that in the forty years since the conductivity measurements of Noyes, no attempt has apparently been made to check or extend the observations (54), in spite of improved available techniques and materials of research. Though these studies did not extend to the critical point, there is at present no apparent reason why such apparatus could not be used.

E. LOSS OF WEIGHT METHOD AND VAPOR-PHASE TRANSPORT

If a substance shows no tendency to deposit back on the original charge when cooled, the solubility may be obtained from the loss of weight of that charge (58), which may be isolated from the rest in a capsule or crucible. This principle was applied (44, 58) to the determination of the solubility of silica. If there is some intermediate state of maximum solubility which must be passed through to reach the temperature and pressure of interest, heating and cooling must be rapid to minimize errors from this source. At one laboratory, at least, the bomb containing the charge is quenched in 1 min. to near room temperature. Silica was observed in most cases to remain in colloidal solution in water; occasionally flakes were observed to separate out (44).

A modification of this method was used for the determination of the solubility of silica in water above its critical point (118). In this apparatus the silica was supported in a crucible near the top of an autoclave, this location minimizing, though not eliminating, error due to flooding by the expanding liquid just below the critical point.

By the use of a similar apparatus, the *transport* of silica through the supercritical fluid phase was observed (68, 115, 116, 117). The silica was here dissolved from one crucible, and absorbed by reaction with zinc oxide in a second crucible.

Similar experiments were carried out with chromium (59b) and molybdenum (59d) compounds, using ferric hydroxide as a "getter" for chromium, and calcium and lead oxides to absorb molybdate ions. In these investigations, flooding by liquid was avoided by using less than the critical volume of solution.

In the course of pneumatolytic experiments using a furnace *inside* the autoclave (57, 74), under conditions encouraging some convection, the gas-phase transport of several refractory substances from the furnace and charge at temperatures above 450°C. was observed. Deposits of sillimanite, rutile, corundum, quartz, and amorphous silica were identified.

It is obvious that measurements of solubility by weight loss cannot be carried out accurately if temperature differences exist in the autoclave, resulting in circulation and the deposition of dissolved material at cold spots, and consequent positive error in solubility values.

F. BALANCING METHOD

One apparatus for establishing the critical points of solutions and rough measurement of solubility at the critical point consisted of a steel tube mounted on knife edges, the whole being heated in an air bath (59c, 59e). When two phases co-existed in the tube, it was unbalanced and tilted, with the denser phase in the lower end. When the contents of the tube, on passing through the critical point, became homogeneous, the tube assumed a horizontal position. When a small amount of solute remained in the solid phase or was precipitated, the tube did not become horizontal, but assumed some intermediate angle. Since the angle of tilt could be fairly accurately measured, the final tilt could be observed, giving the critical temperature of the saturated solution. The critical solubility could be approximated by varying the concentrations of solid in successive tests.

A modification of this method, in which the tube was attached to one arm of an analytical balance (9), has been used for the determination of the critical temperatures of pure liquids, and could no doubt be adapted to solutions.

G. AGITATION

Agitation, or the introduction of any mechanical motion into a high-pressure vessel, is a difficult problem. Leakage along a shaft or chemical action on packing is nearly unavoidable. Filter autoclaves with externally driven stirrers have, however, been developed for use to above 200 atm. pressure and 400°C. (72, 110).

In most hydrothermal experiments the use of stirrers has been avoided. A small charge is ordinarily used, and sufficient time given for equilibrium to occur through diffusion. The filter autoclave of Königsberger and Müller, however, was agitated by simply rocking the entire furnace containing the bomb (60, 61, 62).

The apparatus used at Western Reserve University (11, 11a, 12, 13) for agitation in most high-pressure studies is the magnetic stirrer. This consists of a small piece of iron, commonly sealed into glass, placed inside the high-pressure cell, and activated by the intermittent magnetic field of a solenoid surrounding the cell. This type of stirrer should be usable for other than glass cells. The design of a system can, of course, be much simpler when a magnetic stirrer is used, eliminating leaks due to entering shafts.

H. MATERIALS

Pressures and temperatures encountered in critical studies do not always present the greatest experimental difficulties; the critical points for most stable organic compounds can be observed in sealed borosilicate glass tubes and by other methods (9, 12, 13). Ordinary carbon steel containers could be used, except where ionizable substances are concerned. With water and many other inorganic materials, however, corrosion presents a serious problem.

Most glasses are disintegrated within a few hours by the action of pure water under elevated pressure at 300° to 500°C., although glasses containing large amounts of lead and barium are somewhat more resistant. Borosilicate and soda-lime glasses are very susceptible to such corrosion (71). Although fused quartz is observed (55, 68, 118) to have an appreciable solubility in water at temperatures near the critical point, it has been used successfully (92) for observations on the behavior of aqueous solutions in the critical region. Crystalline quartz, used for electrical insulation (82) at elevated temperatures, was observed to be little affected by long exposure to usually somewhat milder conditions, although hydrated silica under like circumstances was rapidly dissolved.

Several early investigators (73) overcame both the fragility of glass and contamination due to corrosion of metal by sealing the sample in a glass capsule which was enclosed in a steel pressure vessel. In the steel container, outside the glass tube, was placed some of the solvent used in the sample, the effect being to equalize approximately the pressures inside and outside the glass. This principle was used again recently (72) in an intricate filter autoclave.

Ordinary steels are, in general, too easily corroded for use with water in the critical region. Earlier workers used tool steels, soft steels (82), or nickel steels (60, 61, 62), lined with gold, platinum, or platinum–iridium alloy. A silver-plated copper–nickel alloy calorimeter was used (83) in the measurement of thermal properties of water to 270°C. Later work was done with a special chrome–nickel alloy steel (84). Alloys based on 18-8 stainless steel have been used commonly by recent investigators, in particular by Keyes (59) in his well-known studies of the thermodynamic properties of steam. A tendency of pure nickel to creep and crystallize was reported (59).

The problems of gaskets for closures and valve packing have been among the most serious encountered in all but the sealed-tube systems. Ordinary gasket materials are too soft, and usually chemically unsuitable, for most critical determinations. Gold and annealed copper have been used successfully; a gold–copper alloy was, however, discarded (82) in favor of pure gold. At temperatures above 250°C. gold welds under pressure to noble metals, and hence is often not usable as gasket material when the autoclave is lined or plated with such metals (82, 44). This difficulty is eliminated in the recent common use of unlined alloy steel containers.

A metal-to-metal cone joint (6, 15, 59) eliminates altogether the use of gaskets in contact with chemicals. In this design a tight joint is made by a 57–59° cone fitting into a 60° conical seat, leaving nothing but the resistant metal in contact with the charge. This closure requires very precise machining and somewhat more complex design to give a tight seal and to avoid abrasion of the units.

Sampling devices, although they may have a metal-to-metal joint at the valve seat itself, usually require some type of packing on the valve stem. For temperatures not too elevated, ordinary packing materials may suffice for some systems. With water in the critical region, however, none of these is satisfactory. A packing consisting of alternate layers of talc and graphite, precisely cut, was used successfully for aqueous solutions near the critical point (55).

I. MISCELLANEOUS INVESTIGATIONS

A number of studies, while not strictly solubility measurements of the type stressed here, are, however, of fundamental importance and essential to a complete picture of these processes.

Most closely related are investigations concerned with the solubility of water in rock-forming minerals near their melting points, essentially a study of the upper loop of figure 3 (37, 38, 39, 40, 41, 42).

Studies have been carried out relating volume changes due to temperature, pressure, and admixture of various solutes in aqueous systems (34, 35, 36). In connection with this it was noted, for example, that the solubilities of calcium carbonate and sulfate in water were increased 50 per cent by an external pressure of 1000 atm. It is suggested, however, that the reverse may be true for certain magmatic components. Further information has been obtained concerning solubility as a function of temperature and pressure (58).

Of a similar nature is the careful measurement of pressure–volume–temperature relations for salt solutions of different concentrations (6).

VI. Systems Reported in the Literature

A. INTRODUCTION

In this chapter are reviewed the most important systems described in the literature, and those most prominent in the development of the solution above the critical point. The arrangement of systems is alphabetical, the name of the more volatile component coming first. Studies of the electrical properties of solutions in the critical region are reviewed together at the end of the chapter.

B. INDIVIDUAL SYSTEMS

1. Systems with ammonia

The solubilities of a number of organic compounds in ammonia, extending up to the critical point, have been measured by Centnerszwer and others (20, 23). The synthetic method was used, known quantities of solute and ammonia being placed in sealed tubes. In no case were true critical phenomena of the saturated solutions observed. The solutes tested included diphenylmethane, triphenylmethane, naphthalene, anthracene, phenanthrene, α-naphthol, β-naphthol, resorcinol, diphenylamine, α-naphthylamine, β-naphthylamine, benzil, anthraquinone, camphor, dimethyl-*d*-tartaric acid, and urea. The raising of the critical point of ammonia due to dissolved substances was found to be 13.0°C. per mole per cent of solute.

2. System: carbon dioxide–iodine

Cailletet and Colardeau (19) dissolved iodine in liquid carbon dioxide in a sealed tube. When the solution was heated past its critical point, the rose color of dissolved iodine remained in the lower part of the tube. The spectroscope showed the absorption spectrum characteristic of dissolved iodine, not the fluted spectrum of iodine vapor. The interpretation of these data was that although the meniscus had disappeared, the liquid state persisted above the critical point.

The more thorough experiments of Villard (119, 120) led to a more logical and more complete explanation of this system. Upon agitation or long standing of the iodine–carbon dioxide system above its critical point, the iodine coloration became uniform throughout the entire tube. Diffusion of iodine into the vapor phase did not occur in the presence of liquid except with considerable agitation. Any coloring of the vapor phase was in general considerably fainter than that of the liquid, the difference approaching zero as the critical point was approached. Crystals of iodine in the top of the tube dissolved in the vapor, sometimes making the vapor temporarily more highly colored than the liquid phase.

Experiments with iodine in compressed, but unsaturated, carbon dioxide at 20–25°C. showed that an unsaturated vapor may dissolve solids. It is not necessary to postulate liquid dissolved in vapor to explain the solubility of solids in the gas phase.

Spectroscopic examination by Villard of solutions of iodine in both saturated and unsaturated carbon dioxide vapor also showed only the spectrum of dissolved iodine, not of iodine vapor.

3. Systems: carbon dioxide with various solutes

The solubilities of a large number of organic and inorganic substances in carbon dioxide in the region of the critical point were measured by Büchner (16). Many organic compounds, particularly the more volatile, were found to give a complete range of solubilities, eliminating critical phenomena, except of unsaturated solutions. Many other organic compounds, however, and almost all inorganic compounds except gases had more limited solubility, passing through critical points in saturated solution. Organic compounds giving critical points with carbon dioxide included naphthalene, phenanthrene, methyl iodide, *p*-dibromobenzene, borneol, substituted phenols (except nitrophenol), *p*-chloronitrobenzene, *p*-bromonitrobenzene, 1,2,3-dichloronitrobenzene, 1,3,2-dichloronitrobenzene, phthalic anhydride, succinic anhydride, α-naphthylamine, urea, and benzamide. The inorganic compounds tested were not listed. The compounds giving unbroken solubility curves included *p*-dichlorobenzene, acetic acid, camphor, ethanol, carbon disulfide, ethyl ether, pentane, amylene, acetylene, benzene, and *p*-xylene.

The raising of the critical point of carbon dioxide by dissolved substances was found to be 8.8°C. per mole per cent of solute. Comparison of this result with the data of Centnerszwer (20, 23) led to an interesting approximate empirical relationship between the absolute critical temperatures of the pure solvents ammonia, sulfur dioxide, and carbon dioxide, and the changes in these temperatures due to the presence of 1 mole per cent of dissolved material:

$$\Delta T_c = 0.029 T_c$$

4. Systems with carbon disulfide

Hannay and Hogarth (50, 51, 52, 53) observed no deposition of sulfur from solution in carbon disulfide 50°C. above the critical point. Arsenic was also soluble in this region, apparently as a sulfide.

5. System: ethane–hexachloroethane

Holder and Maass (55a) measured the solubility of hexachloroethane in ethane in both liquid and vapor phases, as well as above the critical point of the solution. Solubilities were determined by sampling the solutions in the upper and lower parts of the phosphor bronze bomb used, the material being displaced by mercury, a technique avoiding the disturbance of equilibrium. The proportions of the two components and the partial density of ethane in each phase were obtained by weighing the small sampling bombs, evaporating out the ethane, and reweighing.

It was observed that a difference in the proportions of hexachloroethane in solution in the upper and lower regions of the container, respectively, persisted for 4°C. or more above the previously observed point of disappearance of the meniscus of a saturated solution, the two solubilities approaching each other asymptotically. This was interpreted as a persistence of liquid-like structure over a limited range above the conventional critical point defined by the disappearance of the meniscus.

The partial densities of the ethane in the upper and lower parts of the bomb were found to give curves very similar to the solubility curves, though covering a narrower percentage range of variation. A direct relation between solvent power and partial density of solvent is indicated.

Good precision in solubility measurements in this system is made possible by the solubility of about 14 per cent by weight of hexachloroethane in ethane at the critical point, and in general by the ease of handling and lack of corrosive properties of the substances. The consistency of results and the use of a Podbielniak column in the purification of the ethane, as well as other refinements, suggest that the unique results of Holder and Maass be seriously considered, particularly in view of their support of certain recent theories favoring a critical *region* of finite extent, rather than a critical point (54a, 65a).

6. Systems: ethanol–inorganic compounds

The pioneer studies of Hannay and Hogarth (50, 51, 52, 53) on solutions above the critical point included solutions of several inorganic salts in ethanol. Potassium bromide, potassium iodide, cobalt chloride, and ferric chloride all gave clear gas-phase ethanol solutions far above the critical point. Observations extended to about 380°C., 150° above the critical temperature. Solubility was observed to depend upon both temperature and pressure.

The absorption spectrum of cobalt chloride in ethanol solution was the same at 300°C. as at 15°C., indicating no change in electronic structure on passing into the critical region.

The rather complex behavior of calcium chloride–ethanol solutions above their critical points, including the re-formation of two fluid phases at an elevated temperature, was said to give evidence of compound formation. The later interpretation of Smits (96), however, is that the observed phenomena are typical of a system of the ether–anthraquinone type, having two critical end-points. Smits also believed the potassium bromide and potassium iodide solutions in ethanol to be of this type.

Ramsey (86) heated a solution of potassium iodide in ethanol to above its critical point. Observing that solute crystallized out under these circumstances, and that solution above the critical point failed to dissolve iodide crystals, he concluded that the solubility reported by Hannay and Hogarth was simply solution in hot liquid, the liquid phase persisting above the critical point. This position was strengthened by his critical experiments with presumably pure compounds in sealed tubes. In these experiments the refractive indices in the upper and lower parts of the tubes were found unequal after the disappearance of the meniscus.

Hannay (48) ascribed the results of Ramsay to traces of impurities, particularly water, which, he had observed, lead temporarily to appreciable differences in properties between the upper and lower parts of the contents of a sealed tube.

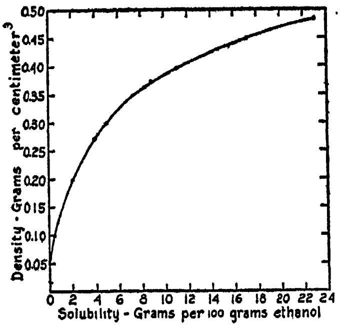


FIG. 7. Solubility of sodium iodide in ethanol above the critical point as a function of solvent density (Tyrer (111)).

Saturation with potassium iodide was found to raise the critical temperature of ethanol by 0.8°C. On the basis of his summed observations, Hannay (49) concluded that solvent power increases with temperature in the single-phase region, at constant specific volume.

Tyrer (111) determined the solubilities of potassium iodide in methanol and sodium iodide in ethanol up to and above the critical temperatures of the saturated solutions. The synthetic method was used, known quantities of salt and alcohol being heated in sealed glass tubes with agitation, and the temperatures of disappearance and reappearance of crystals being noted.

The solubilities were observed to reach a maximum some distance below the critical temperature, and dropped very rapidly in the neighborhood of the critical point. A sharp break occurred at the point where one of the phases disappeared, whether with or without critical phenomena. In the single-phase region, the solubility at constant volume decreased slowly with increased temperature, and was dependent upon the density of the solvent and therefore upon the pressure (figure 7). Solubility in the vapor phase a few degrees below the critical temperature was measured in some cases.

The rapid fall in the solubility of salts in the liquid phase just below the critical temperature was explained by the large dilation of the liquid phase in that region, a corollary of the effect of "solvent concentration" on solubility above the critical point.

7. *Systems: ethanol–organic compounds*

The solution of a chlorophyll extract in ethanol was observed spectroscopically by Hannay and Hogarth (50, 51, 52, 53) above its critical temperature. The spectrum of the solution at 300°C. was the same as at 15°C., indicating that no change in electronic structure occurred on passing through the critical region.

Pictet (85) used a solution of alizarin in ethanol to demonstrate that a solid remains in solution above the critical point. The vapor phase remained colorless until the meniscus disappeared, after which uniform diffusion throughout the tube occurred.

The same result was obtained by Bertrand and Lecarme (7). They observed also the raising of the critical point of ethanol by the dissolved solid, which for a 5 per cent solution of alizarin amounted to 7°C.

8. *System: ethyl ether–anthraquinone*

The ether–anthraquinone system, investigated principally by Smits (95, 96, 97, 102), is the classic example of a system having two critical end-points (figures 2, 3, 4). When ether is heated with excess anthraquinone, the liquid and vapor phases coalesce at critical point P, at 203°C., with 0.15 mole per cent anthraquinone in solution. Upon further heating more anthraquinone is dissolved, and eventually two fluid phases are again formed, this second critical point Q occurring at 247°C. and 13 mole per cent anthraquinone. Upon further heating increased solubilities of anthraquinone occur in the liquid phase, up to the triple point of anthraquinone.

In some of the later experiments (97) pressure and volume, as well as temperature, measurements were made, permitting the drawing of isobars (figure 4) giving concentration of anthraquinone as a function of both temperature and pressure in the region of a single fluid phase.

The phase relations for this type of system were graphically developed, and used to design a space model (96, 102).

9. *Systems with ethyl ether*

The system ethyl ether–ferric chloride was one of those used by Hannay and Hogarth (50, 51, 52, 53) to show visually the persistence of the solution above its critical point. In this state the yellow color remained, diffused uniformly through the tube, but eventually disappeared as reaction with the solvent occurred.

A number of ethereal solutions were investigated by Pictet (85), who measured the raising of the critical temperature by the dissolved solute. His data were apparently not sufficiently precise for the establishment of any general relationship between temperature rise and concentration. The solutes studied in ether solution included borneol, cineole, terpineol, phenol, guaiacol, and iodine.

Smits (96) found saturated solutions in ether of mercuric iodide, silver nitrate, potassium nitrate, cadmium iodide, and alizarin to be of the ether–anthraquinone type, having critical end-points. The ether–anthraquinone system is described separately in this paper. Saturated solutions of *p*-hydroxybenzoic acid and borneol in ether showed no critical phenomena.

10. Systems: hydrogen–alkali metals

Hannay and Hogarth (49, 50, 51), after their initial success in demonstrating the solubility of solids in gases, investigated certain more extreme systems: solutions of alkali metals in hydrogen gas at high pressure. Samples of sodium, potassium, and lithium were placed in high-pressure containers, also containing mercury, not in contact with the alkali metals. Dry hydrogen was then admitted up to pressures of from 300 to 1000 atm. The alkali metal gradually disappeared, and the mercury became pasty, owing to the formation of amalgam. This process occurred in about half an hour at room temperature and within a few seconds at 100°C. The alkali metals were presumably transported through solution in hydrogen to the mercury.

11. Systems: methanol–inorganic salts

Centnerszwer (21) measured the solubilities of several inorganic salts in methanol up to, but not above, the critical points of the solutions by heating known quantities in sealed tubes. The system methanol–potassium iodide was studied with particular care. The raising of the critical point by dissolved salts was found to be proportional to the molality of the salt. Of the many salts investigated, most were found to be very insoluble in methanol, having no appreciable effect on the critical temperature. The salts having measurable solubility at the critical point were sodium chloride, potassium bromide, barium bromide, sodium bromide, mercuric chloride, cadmium chloride, potassium iodide, mercuric bromide, mercuric iodide, and cadmium iodide, in order of increasing solubility at the critical point.

The observations of Tyrer (111) on potassium iodide in methanol are described under "Systems: ethanol–inorganic compounds."

12. Systems with sulfur dioxide

The solubility measurements on organic compounds in ammonia, carried out by Centnerszwer and others, were repeated on the same compounds with sulfur dioxide as solvent, hydroquinone being added to the list (20, 23). Similar results were obtained, no system showing critical phenomena in the presence of free solute. The critical point of sulfur dioxide was raised 11.3°C. per mole per cent of dissolved material.

13. Systems: water–inorganic compounds

The high critical temperature and pressure of water, and its corrosiveness in that region, have discouraged many investigators from the study of aqueous systems, despite their outstanding importance in natural processes. Bertrand and Lecarme (7), however, observed that potassium dichromate remained in solution in water on passing through the critical point, although it deposited if the water all vaporized before the critical temperature was attained. This is in harmony with the low vapor-phase solubility below the critical point noted for alcoholic (111) and other solutions. The dichromate in solution above the critical point diffused throughout the entire tube and was still in solution at 20°C. above the critical point.

Bertrand and Lecarme postulated, as an explanation of their observations, a solution of liquid in gas in equilibrium with a solution of gas in liquid, the two becoming identical at the critical point, and the liquid in gaseous solution persisting for a considerable distance above the critical temperature.

Schröer (92) made a careful study of the raising of the critical temperature by salts in aqueous solution. Solutions of alkali and alkaline earth halides in water at various concentrations were heated to above the critical point in sealed fused-quartz tubes. The behavior of the solutions on heating and cooling was observed. Unambiguous results were obtained only with alkali halides, and only up to 0.684 *M*. Complex phenomena were observed in the case of alkaline earth chlorides, where the solubility curve cuts the critical curve. In some cases two liquid phases were formed.

The dispersion of solute in the gas phase above the critical point was demonstrated by colored solutions of cupric chloride and cobaltous chloride. Prolonged heating of these and other substances led to decomposition.

The raising of the critical point of water by dissolved salts was found to follow the law:

$$\Delta T_c = x^{1/n} + C$$

where n is greater than 1. Schröer gave a theoretical development of this formula, which disagrees more in form than in quantitative evaluation with the linear relation of Centnerszwer (20, 23). Observations were made on aqueous solutions of sodium chloride, potassium chloride, potassium bromide, potassium iodide, rubidium chloride, calcium chloride, strontium chloride, barium chloride, cupric chloride, and cobaltous chloride.

No quantitative observations of solubility in the gas phase were made by Schröer.

Khitarov and others (59b) heated aqueous solutions of chromic acid, sodium chromate, and chromic sulfate to above their critical points in a steel tube. The chromium compound in the gas-phase solution was absorbed in ferric hydroxide in a crucible at the top of the tube. The volume of solution placed in the tube was carefully calculated, to avoid flooding of the crucible by liquid. In all cases a considerable quantity of the chromium compound was transported through the gas phase, to react with the ferric hydroxide. The possibility of pneumatolytic transfer of chromium in any of the forms tested was thus established.

Sodium molybdate solutions were heated in the same apparatus (59d), the crucible at the top of the bomb containing either calcium oxide or lead oxide, to absorb molybdate ion. Although only a trace of molybdenum was transported through the gas phase from pure sodium molybdate, the presence of silica gel greatly enhanced the transfer, to give concentrations of up to 21 per cent molybdic oxide in the crucible after 5 days at 430°C. This transport was greatly diminished by the addition of alkali. These observations agree with the postulated geochemical transport of molybdenum with silica in the form of a heteropoly acid.

In connection with the same project (59a, 59c, 59e) aqueous solutions of a number of compounds and mixtures were heated to above the critical point of water in a silver-plated steel tube balanced on knife edges. The tube was heated in an air bath, its angle of tilt indicating the state of its contents. A horizontal position showed the contents to be a homogeneous gas or gas-phase solution. Presence of solid or liquid phases was indicated by deviations from the horizontal.

When no excess solid phase was present, the critical point of the solution in the tube could be determined directly. In the presence of excess solute the critical point of the saturated solution was observed, and the critical solubility could be estimated from a series of runs, including concentrations above and below that at the observed critical point.

The accuracy of the apparatus was checked by comparison with the observations of Schröer (92). The raising of the critical point of water by a number of inorganic solutes was measured. By comparing the raising of the critical point by a number of acidic substances, notably silica, boric acid, alumina gel, molybdic acid, and chromic acid, with that produced by their alkali salts, it was seen that the effect of a salt was much greater than that of an equal molarity of the corresponding acid. From this it was concluded that alkali ions have an effect on the structure of water, inhibiting gasification. It was not made clear why this effect could not be interpreted in terms of the respective degrees of ionization and association of the several substances, the raising of the critical point being considered as a colligative property.

Solutions containing both sodium metasilicate and boric acid in varying proportions, with constant total molarities of solute, showed a minimum critical point at near-equal mole concentrations, indicating formation of a complex (59a). The same phenomenon was observed for solutions containing molybdic oxide and silica, giving additional support to the heteropoly acid concept (59d).

Solubilities in steam of compounds of importance in boiler chemistry were measured by Spillner (104). The solubilities of sodium chloride, sodium hydroxide, sodium sulfate, and potassium chloride in saturated and superheated steam at various temperatures and pressures were determined. Solubility became appreciable at about 75 kg. per square centimeter pressure. Salts were found nearly equally soluble in saturated and in superheated steam. In general, solubility at a given pressure decreased with temperature increase. The techniques used were the passing of superheated steam over dry solute, and the boiling of solutions of known concentration.

The experiments of Spillner and others were criticized by Straub in his review on turbine blade deposits (106). Straub, in his much more complete and more precise measurements of solubilities in steam, found that the solubilities of sodium chloride, sodium hydroxide, potassium hydroxide, sodium sulfate, and silica in steam varied sharply with both temperature and pressure. Respective solubilities in vapor and liquid phases were seen to approach each other as the critical region was approached. No data were taken close to the critical point. Measurements above the critical temperature were made on superheated steam, that is, at a lower pressure.

Hydrothermal and pneumatolytic syntheses of minerals, many of which involved critical transitions, were reviewed by Morey and Niggli (75) and Morey and Ingerson (73). In the papers reviewed, however, the emphasis was on the products of the reactions, and little information was obtained relative to solubilities in the gas phase.

14. System: water–silica

Because of its outstanding importance, both geologically and from an engineering viewpoint (70, 106), the silica–water system has been studied more frequently and in greater detail than systems of other substances involving critical phenomena or solution in the vapor phase (56a).

In 1930 van Nieuwenberg and Blumendal (115) observed gas-phase transfer of silica. Two crucibles, containing silica and zinc oxide, respectively, were supported near the top of a pressure vessel containing a calculated quantity of water. The apparatus was heated to 400°C. and 300 atm. for several hours. On cooling, the weight of silica was observed to have decreased, the weight of the zinc oxide crucible having correspondingly increased. Zinc silicate (willemite) was detected in the second crucible. The transport of silica through the gas phase, presumably through the medium of solution, was thus demonstrated.

The above experiments were interpreted by Smits (103) in terms of phase study. The silica–water system is of the same type as the ether–anthraquinone system (96, 97) (figure 4), but complicated by the several crystalline modifications of silica. Phase diagrams comparing pressure–temperature–composition relations of the two systems were drawn.

The experiments of van Nieuwenberg and Blumendal were repeated by Morey (68). The experimental conditions were similar, except that the volume of water was carefully chosen to avoid flooding of the crucibles by the dilated liquid just below the critical point. A small loss in weight of silica was observed, presumably due to vapor-phase transport, but of lesser magnitude than in the earlier experiments.

van Nieuwenberg and Blumendal (117) explained this discrepancy by the greater relative quantity of water used in their experiments. The capacity of their container was actually less than the critical volume of the contained water, a result of which was a transitory flooding of the crucibles by liquid during both heating and cooling. That this flooding was merely a source of error rather than the prime vehicle of transfer was established by runs of varying lengths in a later experiment (118). In this experiment, as well as others, the dependence on pressure of solubility in the gas phase was amply demonstrated, explaining the disagreement of Morey's results.

The solubility or volatility of silica with steam having been demonstrated, a quantitative measurement of it as a function of temperature and pressure was made by van Nieuwenberg and van Zon (118). Runs were made essentially as in the earlier experiments, but with only a single crucible, containing quartz. Temperatures were varied from 380° to 420°C., and pressures from 293 to 500 kg. per square centimeter, adjusted by the quantity of water placed in the container.

The solubility of quartz was determined from the loss of weight of the contents of the crucible. Runs of varying lengths established the error due to flooding of the crucible by liquid, which was used as a correction factor, depriving the final results of any great precision. Results (figure 8) are in agreement with the theory of Smits, and correlate well with the similar ether–anthraquinone system.

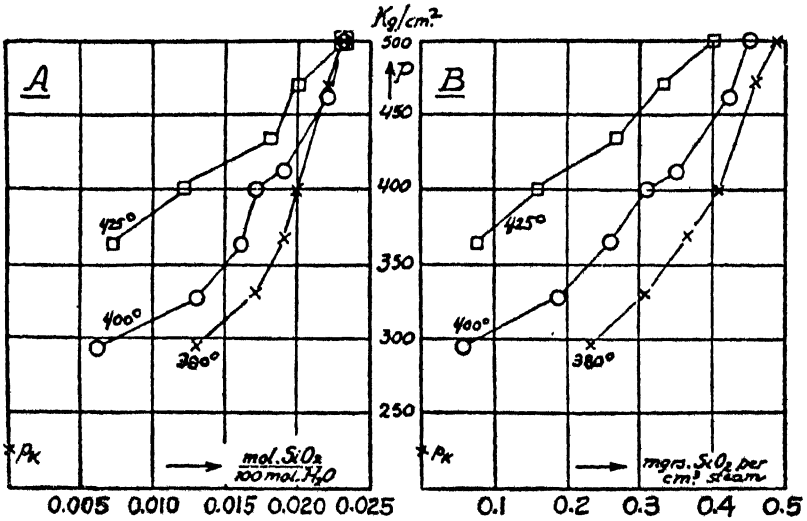


FIG. 8. Solubility of silica in water above the critical point as a function of temperature and pressure (van Nieuwenberg and van Zon (118)).

Ingerson and Morey (57, 70) heated samples of quartz and amorphous silica with approximately the critical volume of water in a hydrothermal bomb. Approximate measurements of solubility at 300°, 400°, and 500°C. were made. At 500°C. 0.5–1.0 per cent of amorphous silica dissolved in the gas phase, and at 400°C. half as much. The solubility of quartz at each temperature was half that of the amorphous silica. The solubility at 300°C. agreed with the 0.17 per cent of Hitchen (55).

Kennedy (58) made a series of solubility measurements on fused quartz along the 300-atm. isobar from 200° to 420°C., by the method of weight differences. Reconnaissance measurements up to 500°C. and 1000 atm. were made. Results suggested that the solubility of silica in steam is an inverse function both of temperature and of specific volume. As the density of the gas phase approaches that of a liquid, the solvent power approaches that of liquid.

Grieg, Merwin, and Shepherd (42) heated pieces of rock, ranging from basalts to obsidians, to 600° to 1200°C. for periods of from 16 hr. to 2 weeks in evacuated silica tubes. Silica evaporated from the tubes, to deposit as fine crystals on platinum foil wrapped around the rock. When no platinum was present, a glassy coating formed on the rock. These phenomena were explained by the dissolving of the silica of the tube by water vapor liberated from the rock, with subsequent deposition (*a*) in the *less soluble* crystalline form, or (*b*) in a solution or compound with the elements of the rock, having a lower vapor pressure of silica than pure fused silica. This hypothesis was proven by later experiments with dry synthetic minerals in the presence or absence of water. The presence of a very small quantity of water greatly increased the quantity of such crystalline or glassy deposits.

Additional evidence of the transfer of solids through vapors was obtained by Morey and Ingerson (74) with a hydrothermal bomb having a furnace inside the bomb. Silica was observed to deposit at colder points of the container and around any leak, where a release of pressure occurred.

In a run of several days' duration, refractory substances were observed to deposit on the refractory furnace core and on platinum leads, the deposit being zoned as the temperature dropped from estimated levels of 1000° to 500°C. The successive deposited materials were sillimanite, rutile, corundum, quartz, and amorphous silica. The steam pressure was 200 atm. The deposited materials were believed to have dissolved in the steam, to be re-deposited at points where their solubilities decreased owing to lowered temperature.

With the recent development of steam turbines operating at pressures of the order of 75 atm., the deposition of silica on turbine blades has become a serious problem. Silica and certain salts are carried from the steam boiler through the superheater in the vapor phase, and deposit on the turbine blades as the temperature and pressure become lower. As mentioned earlier in this paper, a number of studies have been made with the object of preventing this deposition. Although none of these observations extend above the critical point, they are of interest and are related to the problem, since they concern solubilities in the gas phase. The work done on this subject was reviewed by Morey (70) from a phase study viewpoint. A later review by Straub (106) is more concerned with the engineering features of the problem, and presents the results of his extended experiments, as well as their interpretation in terms of the various theories of turbine blade deposits.

The experiments of Straub involved a variety of boilers and bombs, in which saturated or superheated steam was brought in contact with solid silica and solutions containing silica or silicates in known concentration. The steam leaving these containers was condensed and analyzed for silica. The solubility of silica in superheated steam was found to vary with both temperature and pressure (figure 6). The ratio of the solubility of silica in steam to its solubility in the liquid with which it was in contact was found to vary inversely with pH.

C. ELECTRICAL PROPERTIES OF SOLUTIONS

Conductivity measurements of both liquid- and vapor-phase solutions of sodium iodide and sodium bromide in sulfur dioxide were made by Hagenbach

(45, 46) at temperatures extending to above the critical point. The resistance of the liquid solutions increased rapidly up to the critical point, but only a slow increase in resistance occurred above that point. Conductivities of liquid and vapor phases approached each other with rising temperature, becoming identical at the critical point.

The conductivities of solutions of cadmium chloride, calcium chloride, and acetic acid in ethyl ether were measured, but the increasing conductivity of glass made it impossible to continue to the critical region.

Eversheim (30) made conductivity and dielectric constant measurements on ethyl ether, ethyl chloride, and sulfur dioxide, and solutions in these solvents. The conductivity measurements on sulfur dioxide were in complete agreement with those of Hagenbach (45, 46). The dielectric constants were measured by a sealed-in condenser of four concentric cylindrical plates. An abrupt, though not large, change in dielectric constant occurred at the critical point, but little change was observed in the single-phase region at constant volume.

Walden and Centnerszwer (124), in the course of extensive investigations of the properties of sulfur dioxide, observed the conductivity of sulfur dioxide and its solutions to become zero at the critical point. This is, of course, in disagreement with Hagenbach (45, 46). A conclusion from this observation was that the dissociating power of solvents depends upon their *surface tension*, as well as other properties. Electrolytes used in these determinations included hydrogen chloride, quinoline, triamylammonium iodide, benzylammonium chloride, and trimethylsulfine iodide.

Hagenbach (47) explained the discrepancy between his results and those of Walden and Centnerszwer by the greater sensitivity of his apparatus, which was capable of measuring small currents which seemed equal to zero with less refined equipment.

Kraus (63, 64) made a series of very precise determinations of the electrical conductivities of solutions of potassium iodide and ammonium chloride in methanol, and reconnaissance measurements on potassium iodide in ethanol. Both solute concentration and degree of filling of the tubes were varied. Conductivities were measured for some distance above and below the critical point, and were determined in both liquid and vapor phases, where both existed (figures 9 and 10).

Conductivities above the critical point, and of the vapor phase below the critical point, were observed. Though a sudden change in the temperature coefficient of conductivity was observed when liquid was present up to the critical temperature, no abrupt change in the conductivity itself occurred at that point. Unsaturated vapor-phase solutions showed no discontinuity in conductivity at the critical temperature. The conductivity where a single phase existed was found to increase with the density or "concentration" of solvent, and consequently with the pressure. Conductivity fell off slowly with increased temperature above the critical point.

Although most solutions show a maximum in conductivity at some distance below their critical points, the conductivity then monotonically diminishing with further temperature rise, certain solutions show a minimum conductivity in this region (33, 124). Since these solutions have not been carried to the critical temperature, their behavior at that point is unknown.

An electrolytic cell using dry sulfur dioxide as solvent was set up by Hagenbach (47). The E.M.F. values in the liquid- and vapor-phase solutions were found equal at all temperatures, and no singularity of E.M.F. marked the critical point of the solution.

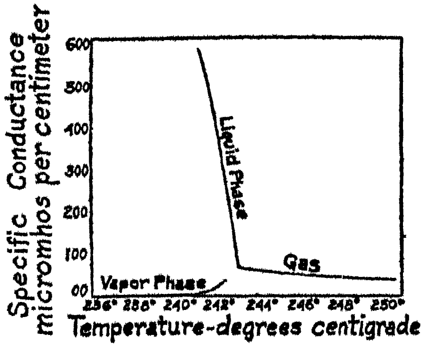


FIG. 9

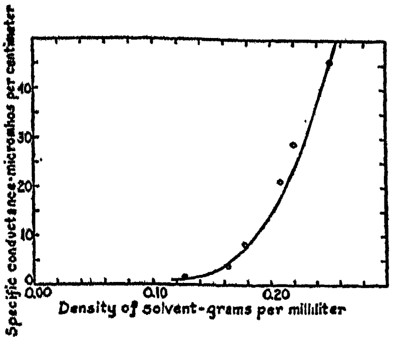


FIG. 10

FIG. 9. Conductance of ammonium chloride in methanol as a function of temperature in the critical region (Kraus (64)).

FIG. 10. Specific conductance as a function of solvent density for 0.00463 *N* potassium iodide in methanol at 245°C. (after Kraus (64)).

VII. CONCLUSION

An effort has been made, in this review, to present to the chemist and the chemical engineer a field of endeavor of which, despite its wide potentialities, comparatively little systematic study has been made. The solubility of solids in vapors and above the critical point has been presented both for its own intrinsic interest, as a challenge in pure research, and from the points of view of its geological and industrial applications. The work in this field on the more important systems reported has been reviewed.

Proper control of the phenomena described here is capable of increasing the over-all efficiency of certain power installations and of prolonging the useful life of turbine and superheater units. The contribution of such study to the understanding of the mechanism of ore deposition may well aid in the more efficient location and exploitation of valuable ore deposits.

Certain synthetic processes may be found to occur more readily under the conditions of high solvent mobility occurring in the critical region (110a). It is possible that further information obtained from these studies may lead to a modification of present concepts of solution, ionization, and association, and may be used as a basis for the growing of large crystals of various substances for optical and other scientific purposes (57a).

It is true that comparatively careful machine work and moderately expensive construction materials are often required for studies at elevated temperatures and pressures, yet the cost can be small as compared with, for example, a recording potentiometer, which is a useful, though not essential, accessory. The principles of the design of such equipment (15, 76, 109) are well understood, and the day is past when the experimenter with high pressure would build his equipment as strongly as he knew how, and then hope for the best. The materials of construction are readily available, and the compounds of greatest interest are among the commonest.

New industrial applications for knowledge of solubilities in the gas phase will probably appear as soon as such knowledge becomes available. The possibilities of this comparatively unexplored field are large, and may be quite unexpected.

VIII. REFERENCES

(1) Allen, E. T.: J. Franklin Inst. **193**, 70 (1922).
(2) Altschul, M.: Z. physik. Chem. **11**, 577 (1893)
(3) American Institute of Physics: *Temperature, Its Measurement and Control in Science and Industry*, pp. 1079–89. Reinhold Publishing Corporation, New York (1941).
(4) Andrews, T.: Trans. Roy. Soc. (London) **A159**, 547 (1869).
(5) Beckmann, E.: Z. physik. Chem. **44**, 161 (1903).
(6) Benedict, M.: J. Geol. **47**, 252–276 (1939).
(7) Bertrand, G., and Lecarme, J.: Compt. rend. **141**, 320 (1905).
(8) Bond, P. A., and Stowe, V. M.: J. Am. Chem. Soc. **53**, 30–4 (1931).
(9) Bond, P. A., and Williams, D. A.: J. Am. Chem. Soc., **53**, 34 (1931).
(10) Booth, H. S.: U. S. patent 2,112,903 (April 5, 1938).
(11) Booth, H. S., and Carter, J. M.: J. Phys. Chem. **34**, 2801–25 (1930).
(11a) Booth, H. S., and Martin, D. R.: Chem. Revs. **33**, 57 (1943).
(11b) Booth, H. S., and Pollard, E. F.: Ind. Eng. Chem. **40**, 1983 (1948).
(11c) Booth, H. S., and Pollard, E. F.: Ind. Eng. Chem. **40**, 1986 (1948).
(11d) Booth, H. S., Pollard, E. F., and Rentschler, M. J.: Ind. Eng. Chem. **40**, 1981 (1948).
(12) Booth, H. S., and Swinehart, C. F.: J. Am. Chem. Soc. **57**, 1337–42 (1935).
(13) Booth, H. S., and Willson, K. S.: J. Am. Chem. Soc. **57**, 2280 (1935).
(14) Bowen, N. L.: *The Evolution of the Igneous Rocks.* Princeton University Press, Princeton (1928).
(15) Bridgman, P. W.: *The Physics of High Pressure.* G. Bell and Sons, Ltd., London (1931).
(16) Büchner, E. H.: Z. physik. Chem. **54**, 665 (1906).
(17) Cagniard de la Tour, C.: Ann. chim. phys. [2] **21**, 127, 178 (1822).
(18) Cailletet, L.: Compt. rend. **90**, 210–11 (1880).
(19) Cailletet, L., and Colardeau, E.: Ann. chim. phys. [6] **18**, 269 (1889).
(20) Centnerszwer, M.: Z. physik. Chem. **46**, 427 (1903).
(21) Centnerszwer, M.: Z. physik Chem. **72**, 431 (1910).
(22) Centnerszwer, M., and Pakalneet, A.: Z. physik. Chem. **55**, 303 (1906).
(23) Centnerszwer, M., and Teletow, I.: Z. Elektrochem. **9**, 799 (1903).
(24) Cleve, K.: Z. Ver. deut. Ing. **84**, 789–95 (1940).
(25) Daubree, A.: Compt. rend. **29**, 227 (1849).
(26) Day, A. L., and Allen, E. T.: *The Volcanic Activity and Hot Springs of Lassen Peak.* Publication No. 360, Carnegie Institution of Washington, Washington, D. C. (1925).
(27) Deville, H. St-C., and Caron: Ann. chim. phys. [4] **5**, 104 (1865).

(28) Deville, H. St-C., and Leblanc, H. E.: Bull. soc. geol. France [2] **13**, 606 (1855–56).

(29) Dewar, J.: Proc. Roy. Soc. (London) **30**, 538 (1880).

(30) Eversheim, P.: Ann. Physik [4] **8**, 539 (1902).

(31) Fenner, C. N.: Am. Inst. Mining Engrs. **1933**, 72.

(32) Foreman, F., and Gruner, J. W.: Econ. Geol. **24**, 825 (1925).

(33) Franklin, E. C., and Kraus, C. A.: Am. Chem. J. **24**, 83 (1900).

(34) Gibson, R. E.: Am. J. Sci. **35A**, 49 (1938).

(35) Gibson, R. E.: *The Nature of Solutions and their Behavior under High Pressures.* Supplementary Publication No. 35, Carnegie Institution of Washington, Washington, D. C. (1938).

(36) Gibson, R. E.: Trans. Am. Geophys. Union **19**, 273 (1938).

(37) Goranson, R. W.: Am. J. Sci. **22**, 481 (1931).

(38) Goranson, R. W.: Trans. Am. Geophys. Union **12**, 163 (1931).

(39) Goranson, R. W.: Am. J. Sci. **23**, 227 (1932).

(40) Goranson, R. W.: Trans. Am. Geophys. Union **17**, 257 (1936).

(41) Goranson, R. W.: Trans. Am. Geophys. Union **18**, 246 (1937).

(42) Goranson, R. W.: Trans. Am. Geophys. Union **19**, 271 (1938).

(43) Grieg, J. W., Merwin, H. E., and Shepherd, E. S.: Am. J. Sci. **25**, 61 (1933).

(44) Gruner, J. W.: Econ. Geol. **25**, 697 (1930).

(45) Hagenbach, A.: Physik. Z. **1**, 481 (1900).

(46) Hagenbach, A.: Ann. Physik [4] **5**, 276 (1901).

(47) Hagenbach, A.: Ann. Physik [4] **8**, 568 (1902).

(48) Hannay, J. B.: Proc. Roy. Soc. (London) **30**, 478 (1880).

(49) Hannay, J. B.: Proc. Roy. Soc. (London) **30**, 484 (1880).

(50) Hannay, J. B., and Hogarth, J.: Proc. Roy. Soc. (London) **29**, 324 (1879).

(51) Hannay, J. B., and Hogarth, J.: Proc. Roy. Soc. (London) **30**, 178 (1880).

(52) Hannay, J. B., and Hogarth, J.: Chem. News **40**, 256 (1880).

(53) Hannay, J. B., and Hogarth, J.: Chem. News **41**, 103 (1880).

(54) Harned, H. S., and Owen, B. B.: *The Physical Chemistry of Electrolytic Solutions*, p. 492. Reinhold Publishing Corporation, New York (1943).

(54a) Harrison, S. F., and Mayer, J. E.: J. Chem. Phys. **6**, 87 (1938).

(55) Hitchen, C. S.: Bull. Inst. Mining Met. **1**, 364 (1935).

(55a) Holder, C. H., and Maass, O.: Can. J. Research **18B**, 293 (1940).

(56) Ingerson, E.: Econ. Geol. **29**, 454 (1934).

(56a) Ingerson, E.: Econ. Geol. **42**, 560 (1947).

(57) Ingerson, E., and Morey, G. W.: Trans. Am. Geophys. Union **22**, 504 (1941).

(57a) Ingerson, E., and Tuttle, O. F.: Am. J. Sci. **245**, 313–19 (1947).

(58) Kennedy, G. C.: Econ. Geol. **39**, 25 (1944).

(59) Keyes, F. G.: Proc. Am. Acad. Arts Sci. **68**, 505 (1933).

(59a) Khitarov, N. I.: Soviet Geol. **9**, No. 7, 71 (1939).

(59b) Khitarov, N. I., and Ivanov, L. A.: Soviet Geol. **6**, 1098 (1936).

(59c) Khitarov, N. I., and Ivanov, L. A.: Zentr. Mineral. Geol. **A1936**, 46.

(59d) Khitarov, N. I., and Ivanov, L. A.: Compt. rend. acad. sci. U.R.S.S. **27**, 694 (1940).

(59e) Khitarov, N. I., Ivanov, L. A., and Rotman, L. E.: Soviet Geol. **9**, No. 2, 98 (1939).

(60) Königsberger, J., and Müller, J. W.: Zentr. Mineral. Geol. **A1906**, 339.

(61) Königsberger, J., and Müller, J. W.: Chem. Zentr. **1906**, II, 268.

(62) Königsberger, J., and Müller, J. W.: Zentr. Mineral. Geol. **A1912**, 413.

(63) Kraus, C. A.: Phys. Rev. **18**, 40, 89 (1904).

(64) Kraus, C. A.: *The Properties of Electrically Conducting Systems*, pp. 167–75. The Chemical Catalog Company, Inc., New York (1922).

(65) Lupberger, E.: Stahl u. Eisen **53**, 1080 (1933).

(65a) Maass, O., and Geddes, A. L.: Trans. Roy. Soc. (London) **A236**, 303 (1937).

(66) Mellor, J. W.: *A Comprehensive Treatise on Inorganic and Theoretical Chemistry*, Vol. II, p. 327. Longmans, Green and Company, London (1922).

(67) Morey, G. W.: J. Wash. Acad. Sci. **12**, 219 (1922).
(68) Morey, G. W.: Trans. Am. Geophys. Union **13**, 269 (1932).
(69) Morey, G. W.: *The Properties of Glass*, p. 103. Reinhold Publishing Corporation, New York (1938).
(70) Morey, G. W.: Proc. Am. Soc. Testing Materials **42**, 980–1022 (1942).
(71) Morey, G. W., and Bowen, N. L.: J. Soc. Glass Tech. **11**, 97 (1927).
(72) Morey, G. W., and Burlew, J. S.: Am. J. Sci. **35A**, 193 (1938).
(73) Morey, G. W., and Ingerson, E.: Econ. Geol. **32**, 607–760 (1937).
(74) Morey, G. W., and Ingerson, E.: Am. J. Sci. **35A**, 217 (1938).
(75) Morey, G. W., and Niggli, P.: J. Am. Chem. Soc. **35**, 1086 (1913).
(76) Newitt, D. M.: *The Design of High-Pressure Plant and the Properties of Fluids at High Pressures*. The Clarendon Press, Oxford (1940).
(77) Niggli, P.: *Das Magma und seine Produkte*. Akademisches Verlagsgesellschaft, Leipzig (1937).
(77a) Niggli, P.: *Ore Deposits of Magmatic Origin*. T. Murby and Company, London; D. Van Nostrand Company, New York (1929).
(78) Niggli, P.: Z. anorg. allgem. Chem. **75**, 161 (1912).
(79) Niggli, P.: Z. anorg. allgem. Chem. **77**, 321 (1912).
(80) Niggli, P.: Zentr. Mineral. Geol. **A1912**, 321.
(81) Norton, F. H.: Am. Mineral. **22**, 1 (1937).
(82) Noyes, A. A., *et al.*: *The Electrical Conductivity of Aqueous Solutions*. Publication No. 63, Carnegie Institution of Washington, Washington, D. C. (1907).
(83) Osborne, N. S., Stimson, H. F., and Fiock, E. F.: Bur. Standards J. Research **5**, 411 (1930).
(84) Osborne, N. S., Stimson, H. F., Fiock, E. F., and Ginnings, D. C.: Bur. Standards J. Research **10**, 155 (1933).
(85) Pictet, R.: Compt. rend. **120**, 64 (1895).
(86) Ramsay, W.: Proc. Roy. Soc. (London) **30**, 323 (1880).
(87) Reilly, J., and Rae, W. N.: *Physico-Chemical Methods*, pp. 598–9. D. Van Nostrand Company, New York (1925).
(88) Roozeboom, H. W. B.: *Die heterogenen Gleichgewichte vom Standpunkte der Phasenlehre*. Friedrich Vieweg und Sohn, Braunschweig (1901–1918).
(89) Schroeder, W. C., Berk, A. A., and Gabriel, A.: J. Am. Chem. Soc. **58**, 843 (1936).
(90) Schroeder, W. C., Berk, A. A., and Gabriel, A.: J. Am. Chem. Soc. **59**, 1783 (1937).
(91) Schroeder, W. C., Gabriel, A., and Partridge, E. P.: J. Am. Chem. Soc. **57**, 1539 (1935).
(92) Schröer, E.: Z. physik. Chem. **129**, 79 (1927).
(93) Shepherd, E. S.: Bull. Hawaiian Volcano Observatory **9**, 83 (1921).
(94) Shepherd, E. S.: Am. J. Sci. **35A**, 311 (1938).
(95) Smits, A.: Z. Elektrochem. **9**, 663 (1903).
(96) Smits, A.: Z. physik. Chem. **51**, 193 (1905).
(97) Smits, A.: Z. physik. Chem. **52**, 587 (1905).
(98) Smits, A.: Z. physik. Chem. **54**, 498 (1906).
(99) Smits, A.: Z. physik. Chem. **54**, 513 (1906).
(100) Smits, A.: Z. physik. Chem. **67**, 454 (1909).
(101) Smits, A.: Z. physik. Chem. **67**, 464 (1909).
(102) Smits, A.: Z. physik. Chem. **76**, 445 (1911).
(103) Smits, A.: Rec. trav. chim. **49**, 962 (1930).
(104) Spillner, F.: Chem. Fabrik **13**, 405 (1940).
(105) Splittgerber, A.: Vom Wasser **15**, 271 (1940).
(106) Straub, F. G.: Univ. of Illinois Bull., Eng. Expt. Sta. Bull. Ser., No. **364** (1946).
(107) Straub, F. G.: Trans. Am. Soc. Mech. Engrs. **1932**, F. S. P. -54-17, p. 221.
(108) Tilden, W. A., and Shenstone, W. A.: Trans. Roy. Soc. (London) **A175**, 23 (1885).

(109) Tongue, H.: *The Design and Construction of High Pressure Chemical Plant.* D. Van Nostrand Company, Inc., New York (1934).
(110) Tongue, H.: Trans. Inst. Chem. Engrs. **8**, 81 (1930).
(110a) Toriumi, T., *et al.*: J. Soc. Chem. Ind. Japan **49**, 1–2 (1946).
(111) Tyrer, D.: J. Chem. Soc. **97**, 621 (1910).
(112) van der Waals, J. D.: *Die Continüität des gasförmigen und flüssigen Zustandes.* Johann Ambrosius Barth, Leipzig (1881).
(113) van der Waals, J. D., and Kohnstamm, P.: *Lebrbuch der Thermodynamik.* Johann Ambrosius Barth, Leipzig (1912).
(114) Reference 113, Vol. 2, pp. 552ff.
(115) van Nieuwenberg, C. J., and Blumendal, H. B.: Rec. trav. chim. **49**, 857 (1930).
(116) van Nieuwenberg, C. J., and Blumendal, H. B.: Rec. trav. chim. **50**, 129, 989 (1931).
(117) van Nieuwenberg, C. J., and Blumendal, H. B.: Rec. trav. chim. **53**, 476 (1934).
(118) van Nieuwenberg, C. J., and van Zon, P. M.: Rec. trav. chim. **54**, 129 (1935).
(119) Villard, P.: Compt. rend. **120**, 182 (1895).
(120) Villard, P.: Ann. chim. phys. [7] **10**, 409 (1897).
(121) Vogt, J. H. L.: J. Geol. **30**, 659 (1922).
(122) Waldeck, W. F., Lynn, G., and Hill, A. E.: J. Am. Chem. Soc. **54**, 928 (1932).
(123) Waldeck, W. F., Lynn, G., and Hill, A. E.: J. Am. Chem. Soc. **56**, 43 (1934).
(124) Walden, P., and Centnerszwer, M.: Z. physik. Chem. **39**, 513 (1902).
(125) Winkler, H. G. F.: Am. Mineral. **32**, 131 (1947).
(126) Zies, E. G.: Natl. Geographic Soc., Contributed Tech. Papers, Katmai Ser. **3**, **159** (1924).
(127) Zies, E. G.: Natl. Geographic Soc., Contributed Tech. Papers, Katmai Ser. **4** (1929).
(128) Zies, E. G.: Chem. Revs. **23**, 47 (1938).
(129) Zies, E. G., and Allen, E. T.: Natl. Geographic Soc., Contributed Tech. Papers, Katmai Ser. **2** (1923).